Requirements of a habitat specialist in Swiss mountain forests – an assessment of forest structure and composition using laser remote sensing and field data

Zellweger, Florian ; Morsdorf, Felix ; Braunisch, Veronika ; Bollmann, Kurt

Abstract: Species richness in forest ecosystems largely depends on habitat structure and composition. These attributes can be assessed in field surveys, however, such data often lacks in spatial extent. Remote sensing technologies such as light detection and ranging (LiDAR) provide alternative tools to quantify structural elements across relatively broad areas at a fine resolution. To study the habitat requirements of hazel grouse (Bonasa bonasia), an indicator species of structurally rich forest stands, we assessed the structure and composition of Swiss mountain forests over three biogeographical regions. We designed a sample based field survey of forest structure and composition and a LiDAR based assessment of vertical and horizontal forest structures using a nationwide LiDAR dataset with a mean point density of 1.4 m2. The dependent variable consisted of hazel grouse presence/absence data at a resolution of 1 km2. Species distribution models were computed for both variable sets separately and in combination, using boosted regression trees, a statistical machine learning technique. Model performance assessment based on explained deviance and AUC showed that the combined model performed best, with over 55% explained deviance in the observed data, followed by the field and LiDAR models. The field model revealed that hazel grouse favored evenly distributed, rich ground vegetation, optimally with a substantial portion of bilberry (Vaccinium myrtillus). The abundance of tall rowans (Sorbus aucuparia), basal branched trees and a high percentage of resource trees in the shrub layer were found to be further essential habitat elements. LiDAR was powerful in detecting important structural features, whereby the horizontal forest structure explained more of the deviance than the vertical forest structure. The most influential LiDAR variable was a measure of canopy height heterogeneity. Apart from indicating structurally rich forest stands, it probably also served as a proxy of compositional aspects such as the abundance of light demanding resource trees and shrubs or of a well developed ground vegetation. To support habitat management, we derived variable thresholds at a relevant spatial scale (1 km2) for forest management. Our study showed that LiDAR provides adequate means to assess structural habitat elements area-wide, thus overcoming the difficulties associated with sample based field assessments. The best model fit, however, was obtained by combining LiDAR variables with compositional variables from the field survey. Hence, we successfully bridged the gap between different ecologically relevant scales, such as habitat configuration and structure at the regional scale and the abundance of habitat elements at the local scale. The methods applied in this study can also be used to identify hotspots of forest structural richness, a matter of interest in the light of emerging attempts to conserve biodiversity in forests.

Originally published at:
Zellweger, Florian; Morsdorf, Felix; Braunisch, Veronika; Bollmann, Kurt (2012). Requirements of a habitat specialist in Swiss mountain forests – an assessment of forest structure and composition using laser remote sensing and field data. In: SilviLaser 2012, Vancouver, 16 September 2012 - 19 September 2012. SilviLaser, online.

Silvilaser2012.com

Full Proceedings

Full Papers

The estimation of forest inventory parameters from small-footprint waveform and discrete return airborne LiDAR data

Sumnall, M. J.[1], Hill, R. A.[2], and Hinsley, S. A.[3]

[1]School of Applied Science, Bournemouth University, Talbot Campus, Poole, Dorset, BH12 5BB, UK. msumnall@bournemouth.ac.uk
[2]School of Applied Science, Bournemouth University, Talbot Campus, Poole, Dorset, BH12 5BB, UK. rhill@bournemouth.ac.uk
[3]Centre for Ecology & Hydrology, Wallingford, Oxfordshire, OX10 8BB, UK. sahi@ceh.ac.uk

Paper Number: SL2012-020

Abstract

The quantification of forest structure is important for a variety of disciplines, including ecology and management, as vegetation structure is related to a wide variety of ecosystem processes. This research investigates the estimation of forest structural variables from small-footprint airborne LiDAR capturing both discrete return (DR) and full waveform (FW) data. The study site is in southern England and contains a variety of forest structural types. The DR and FW LiDAR data were acquired simultaneously using a Leica ALS50-II system, under both leaf-on and leaf-off conditions. The DR data had up to four discrete returns per laser pulse, with an average of ~4 pulses (i.e. first/only returns) per sq. metre. The FW data were acquired at a rate of ~2.5 pulses (i.e. full waveform) per sq. metre. Point data were generated from each waveform through Gaussian decomposition, yielding 1-11 returns per pulse (mode = 5). Metrics derived from each return point included: x-, y- and z-coordinates, plus intensity (DR only), or amplitude and width (FW only). Field data collection was conducted in 21 field plots (30 x 30m in size), located across a stratification of forest structural properties within the study site. A number of forest inventory metrics were recorded per plot, including: (i) tree height; (ii) tree stem count; (iii) tree diameter at breast height (DBH); (iv) tree crown base height; (v) canopy cover; (vi) standing deadwood (snag) volume; and (vii) seedling number. This dataset formed the basis for 'training' subsequent statistical models. For each field plot area several hundred metrics were extracted from both the DR and FW LiDAR data, using both individual tree crown and area-based approaches. These LiDAR metrics were used for direct comparison with field data and in least-squares stepwise linear regression analysis to model forest structure. Compared with field data, dominant height and canopy cover were derived from the airborne LiDAR data with an RMSE of 1.31m and 10% (FW) and 1.32m and 13% (DR) respectively. Of the ten field variables modelled with regression analysis, five were modelled with a higher R^2 value using FW LiDAR data, two were modelled with a higher R^2 value using DR LiDAR data, and three were modelled with very similar R^2 values (i.e. within 0.05) using both LiDAR datasets. Although further analysis and validation is still required, these results suggest that full waveform LiDAR data describe the structural components of woodland to an equal (or better) level of accuracy than discrete return data, in spite of a considerably lower spatial sampling rate due to the waveform digitisation process.

1. Introduction

The quantification of forest structure is important for a variety of disciplines, including ecology and management, as vegetation structure is related to a wide variety of ecosystem processes. Due to the spatially heterogeneous nature of landscapes, woodland inventory can be applied at a variety of scales, from the individual tree to the landscape-scale. Airborne LiDAR systems can

capture information which may be used to estimate forest structural metrics over wide areas (Anderson *et al* 2008; Coops *et al* 2004), and can provide a means to characterise woodlands in a manner previously impossible through fieldwork alone. LiDAR data can be used to describe vegetation structure and to a certain extent to derive information on the height and density of different elements of forest strata. Excellent recent reviews of the application of LiDAR data for forest assessment have been provided by Hudak *et al* (2009) and Wulder *et al* (2012).

Most commercial LiDAR systems used in forest inventory over the past decade deliver discrete return (DR) point data. Due to limitations in the electronics of many airborne laser scanning systems, only surfaces that are sufficiently spaced apart (typically >2.5m) can be distinguished in separate returns. Recently full-waveform (FW) LiDAR has become available which records the whole of the returned laser waveform, potentially allowing for more detailed processing (Persson *et al* 2005) and a more advanced understanding of the factors influencing the return signal over forested areas (Wagner *et al* 2008; Lin *et al* 2010). The potential of small-footprint FW LiDAR systems has already been investigated for tree species classification (Heinzel and Koch 2011; Vaughn *et al* 2011), single tree segmentation (Reitberger *et al* 2009) and above ground biomass estimation (Kronseder *et al* 2012). Pirotti (2011) reviewed DR and FW LiDAR systems for forest inventory based on the published scientific literature. Whilst an extremely useful overview, this did not draw conclusions from directly comparable systems or experimental set-ups. To date then, few studies have approached the issue of directly comparing the capabilities of small-footprint FW and DR LiDAR data for forest inventory. Chauve *et al* (2009) compared the size of point clouds and the resulting quality of Digital Terrain Models (DTMs) and Canopy Height Models (CHMs) from DR and FW LiDAR; whilst Lindberg *et al* (2012) compared the ability of small-footprint DR and FW LiDAR data for forest volume estimation, concluding that FW data characterised volumes to a somewhat better degree than DR data.

This paper compares the estimation of common forest structural variables from small-footprint airborne LiDAR which simultaneously captured both discrete return (DR) and full waveform (FW) data. The study area in southern England contains a variety of forest structural types, including managed plantation, semi-ancient coniferous and deciduous woodland.

2. Field Site

The study site is located within the New Forest National Park in southern England. The New Forest lies to the west of Southampton, and covers 37,677 hectares (Forestry-Commission 2011). The New Forest is recognised as being of international importance to nature conservation for its extensive tracts of semi-natural vegetation and ecologically valuable habitats. This research is focused on a 3x6 km area with the central coordinates of 50° 49’ 33” N, 1° 30’ 3” W. This area includes the Frame-Heath, Hawkhill, Denny, Denny-Lodge and Stockley Incolsures. These woodlands are actively managed by the Forestry Commission and are typically fenced and subject to felling. In addition Tantany, Frame, and Denny Woods are semi-ancient woodland, which are unenclosed and are not subject to felling operations.

The study area contains several types of semi-natural and plantation coniferous and deciduous forests. Broadleaf species include: Oaks (*Quercus robur* and *Quercus petrea*), Beech (*Fagus sylvatica*) and Holly (*Ilex aquifolium*), while coniferous species include: Corsican Pine (*Pinus nigra* var maritime), Scots Pine (*Pinus sylvestris*), Douglas Fir (*Pseudotsuga menziesii*) and Norway Spruce (*Picea abies*) (Koukoulas and Blackburn 2004; Newton *et al* 2010). This array of forest types within close proximity of each other, presents a wide range of available structural variables of interest, such as canopy gaps and the presence of deadwood.

3. Data

Ground survey work was completed between June and September 2010, using 21 square plots located throughout Frame-Heath and Hawkhill Inclosues, in addition to Frame and Tantany Woods, as indicated in Figure 1. Plot locations were established though a random stratification of a Normalised Difference Vegetation Index (NDVI) image, produced from 2007 CASI-2 data. According to Gamon *et al* (1995) and Kalacska *et al* (2004), NDVI values correlate with different forest canopy structural types. The plot size was established as 30m x 30m, with a 10m x 10m subplot located in the south-west corner. A differential GPS (dGPS) system, the Leica GPS 500 (Leica Geosystems) and a Sokkia 6F total station (Sokkia Topcon Co. Ltd.) were then utilised to reference each of the plot corners beneath the canopy, recording each in British National Grid (BNG) coordinates with an accuracy of better than 0.5m. Table 1 summarises the field metrics recorded. In addition, various extra measures were calculated from the field plot data, e.g. tree diameter at breast height (DBH) was derived from the girth measurements. For validation purposes, 20 additional field plots were surveyed in summer 2012. These newer plot locations were identified again by a random stratification of the NDVI image, and recorded the same information as the previous work. The validation plots were located in Denny, Denny-Lodge and Stockly Inclosures, in addition to Denny Wood.

Table 1: Field plot attributes

Tree species metrics
Leaf Area Index, Canopy Openness.
Estimate of Canopy Layers
Tree X- Y- Coordinates, Tree Species, Girth (m) (circumference at 1.3m height), Tree Height (m), Height of First Live Branch (m) (canopy base), Tree Crown Horizontal Extent (m) and assessment of Canopy Health: [1-4]. [For trees with circumference >1.3m]
Number of Saplings and Species (height > 1.3m and circumference <8cm)
Number of Seedlings and Species (height < 1.3m)
Ground vegetation metrics
Species, Height, and Horizontal Extent of Shrub Vegetation (if >1m height)
Vascular Plants Percentage Cover, Estimate of Percentage Bare Soil
Estimate of Percentage Cover of Moss on Trees and Ground
Deadwood metrics
Height of Snags (m), Girth of Snag (m), Snag Decay Class: [Light (1) - Heavy (3)]
Fallen Tree Length (m), Fallen Tree Girth (m), Fallen Tree Decay Class: [Light (1) - Heavy (3)].
Fallen Branch Length (m), Fallen Branch Girth (m), Fallen Branch Decay Class: [Light (1) - Heavy (3)]

Airborne LiDAR data were acquired using a Leica ALS50-II system on two dates, 8th April and 6th July 2010, in order to collect data under leaf-off and leaf-on conditions respectively. Both discrete return (DR) and full-waveform (FW) small-footprint LiDAR data were captured simultaneously using the same scanner for both dates. The data capture specifications are summarised in Table 2. The DR data had up to four discrete returns per laser pulse, with an average of 4 pulses (i.e. first/only returns) per sq. metre. The FW data were acquired at an average rate of 1 pulse (i.e. full waveform) per sq. metre. For the DR data, x-, y- and z-coordinates, intensity, and return number were supplied for the first, intermediate, and last significant return per pulse. For the FW data, 256 return signal amplitude values (sampled every two nanoseconds for the April data and every one nanosecond for the July data) were supplied for each pulse.

Table 2: LiDAR data acquisition specifications

	Discrete Return LiDAR (leaf-off)	**Discrete Return LiDAR (leaf-on)**	**Full Waveform LiDAR (leaf-off)**	**Full Waveform LiDAR (leaf-on)**
Flying altitude (m)	1595	1602	1595	1602
Swath width (m)	551.2	556.0	551.2	556.0
Scan half angle	9.80°	9.84°	9.80°	9.84°
Beam divergence (mr)	0.22	0.22	0.22	0.22
Ground sample rate	~4 pulse returns per m^2	~4 pulse returns per m^2	~2.5 pulse returns per m^2	~2.5 pulse returns per m^2

4. LiDAR data processing and analysis

The DR LiDAR point cloud data required filtering to separate the ground and vegetation hits so that ground elevation could be determined and used to derive the above-ground height of the vegetation within the study area. To that end, the RSCLAS Tools (Armston 2012) software was used which employs a progressive morphological filter, as outlined in Zhang *et al* (2003). Ground elevation values were then removed from the DR LiDAR dataset to yield vegetation height. 30 x 30m polygons were created corresponding to the field plot locations and area statistics were produced for each plot including: maximum, minimum, mean, skewness, kurtosis, *etc*, as in Falkowski *et al* (2009) and Hudak *et al* (2008). These LiDAR metrics were extracted separately for both the leaf-on and leaf-off datasets, and for the points separated into ground and vegetation returns or all returns combined. This totalled 172 metrics from the DR LiDAR data for each 30 x 30m field plot area.

The FW LiDAR data required a number of additional steps in order to derive a point cloud from the recorded waveforms. The FW data were processed utilising Gaussian decomposition (Wagner *et al* 2006) in SPDlib software (http://sourceforge.net/projects/spdlib/). This yielded between 1 and 11 returns per pulse (mode = 5), supplying x-, y- and z- coordinates, amplitude and width per pulse (Bunting *et al*.2011a; 2011b). SPDlib also provided tools for vegetation filtering and normalisation, again using the progressive morphological filter as outlined in Zhang *et al* (2003). As with the DR data, area-based LiDAR metrics were extracted for the field plot areas separately for both the leaf-on and leaf-off datasets, and for the derived points separated into ground and vegetation returns or all returns combined. This totalled 270 metrics from the FW LiDAR data for each 30 x 30m field plot area. The area-based statistics extracted from both DR and FW LiDAR data are summarised in Table 3.

In addition, TIFFS software (http://www.globalidar.com/) was used to delineate individual tree crowns (ITC) in the overstorey using leaf-on LiDAR data (DR and FW separately) and subsequently to calculate their 3D position, height (m), crown horizontal area (m^2) and crown volume (m^3). The extraction technique in TIFFS software is based upon that presented in Chen *et al* (2007). This resulted in a GIS database of individual tree locations and other attributes. These ITC metrics were then summarised and/or totalled from the individual tree to the plot level (30 x 30m), yielding a single summary value for the number of trees within each 30 x 30m area, average tree height, the mean/total horizontal extent and the mean/total volume of the tree crowns. The average and standard deviation of ITC spacing was also estimated and summarised at the 30 x 30m level using R software (http://cran.r-project.org/).

For each field plot, hundreds of metrics were extracted from both the DR and FW LiDAR data, using both individual tree crown and area-based approaches. These LiDAR metrics were used for direct comparison with field data and in least-squares stepwise linear regression to model forest structure. This approach was utilised to test every possible permutation of LiDAR metrics in order to estimate the plot-level forest variables.

Table 3: Summary of extracted LiDAR metrics for both DR and FW data (including variable codes*)

ITC metrics	**Normalised Height:**	**Intensity/pulse amplitude:**	**Pulse width:**
Mean tree heights *(ITC_ht)*	-	Minimum *(min_amp)*	Minimum *(min_wid)*
Mean crown area (m^2) *(ITC_mean_area)*	Maximum *(max_ht)*	Maximum *(max_amp)*	Maximum *(max_wid)*
Total crown area (m^2) *(ITC_tot_area)*	Mean *(mean_ht)*	Mean *(mean_amp)*	Mean *(mean_wid)*
Mean Geometric Canopy Volume (m^3) *(ITC_mean_vol)*	Median *(med_ht)*	Median *(med_amp)*	Median *(med_wid)*
Total Geometric Canopy Volume (m^3) *(ITC_tot_vol)*	Standard Deviation *(STD_ht)*	Standard Deviation *(STD_amp)*	Standard Deviation *(STD_wid)*
Average distance between trees (m) *(ITC_dist)*	Variance *(var_ht)*	Variance *(var_amp)*	Variance *(var_wid)*
Standard deviation of tree distance (m) *(ITC_STD)*	Absolute deviation *(AD_ht)*	Absolute deviation *(AD_amp)*	Absolute deviation *(AD_wid)*
Number of tree crowns in plot *(ITC_tree_no)*	Skewness *(skw_ht)*	Skewness *(skw_amp)*	Skewness *(skw_wid)*
-	Kurtosis *(kur_ht)*	Kurtosis *(kur_amp)*	Kurtosis *(kur_wid)*
-	Percentiles at 5% intervals *(p5,p10…p90,p95)*	Percentiles at 5% intervals *(amp5,amp10…amp95)*	Percentiles at 5% intervals *(wd5,wd10…wd90,wd95)*
DTM metrics	Canopy depth (m) *(can_depth)*	-	-
Ground slope *(Dem_slope)*	Canopy cover *(can_cov)*	-	-
Ground roughness *(Dem_rough)*	Largest gap between canopy layers (m) *(max_vert_gap)*	-	-

*An additional suffix is added to each code (apart from the ITC codes) to show leaf-off (.._loff) or leaf-on (.._lon) and whether the data have been separated into ground returns (.._g) and vegetation returns (.._v) or kept combined (.._all).

5. Initial results

Compared with field data, dominant height and canopy cover were derived from the airborne LiDAR data with an RMSE of 1.31m and 10% (FW) and 1.32m and 13% (DR) respectively. For the modelling, only the results from the least-squares stepwise linear regression for ten field measures are reported here (Table 4). It should be noted that all the models shown in Table 4 and all the individual LiDAR metrics forming an input to the models were significant at $p < 0.05$. Eleven of the models shown in Table 4 had an R^2 value >0.80, and were produced from a mixture of DR and FW data. This included: the number of trees per plot (FW), canopy base height (DR, FW), crown area (FW), mean DBH (FW), standing deadwood volume (DR, FW), the number of saplings per plot (DR, FW), the number of seedlings per plot (DR), and the percentage of the plot with no ground cover (FW). This only left the two forest variables of average tree spacing and

downed deadwood volume for which neither DR nor FW data were able to generate models with an R^2 >0.80. Note that a statistically significant model could not be produced for downed deadwood using the DR LiDAR data.

Of the ten field variables examined, five were modelled with a higher R^2 value using FW LiDAR data (the number of trees per plot, crown area, mean DBH per plot, downed deadwood volume, and the percentage of the plot with no ground cover), three were modelled with a higher R^2 value using DR LiDAR data (average tree spacing, standing deadwood volume, and the number of seedlings per plot), and two were modelled with very similar R^2 values (i.e. within 0.05) using both DR and FW LiDAR data (canopy base height, and the number of saplings).

Table 4 – A sample of the regression results from regression analysis

	DR		FW	
Field Measure	**R^2**	**Variables in model *[3]**	**R^2**	**Variables in model *[3]**
No. Trees	0.695	"ITC_dist" + "kur_ht_lon_all"	0.803	"ITC_tree_no" + "kur_ht_loff_all" + "kur_ht_lon_v"
Tree Spacing (m)	0.640	"ITC_dist" + "can_depth_lon_all"	0.539	"ITC_dist"
Canopy Base Height (m)	0.896	"var_ht_lon_all" + "ITC_dist" + "ITC_mean_vol"	0.928	"max_ver_gap_loff_all" + "ITC_dist" + "ITC_tot_vol"
Crown Area (m^2)	0.748	"ske_ht_lon_all" + "ITC_dist" + "Ske_ht_loff_all"	0.877	"ITC_mean_area" + "kur_ht_lon_v" + "ITC_STD"
Mean DBH (m)	0.663	"ITC_dist" + "ITC_mean_vol"	0.911	"ITC_mean_area" + "skw_ht_lon_v" + "AD_amp_loff_all" + "kur_wid_lon_v"
SDW*[1] Volume (m^3)	0.920	"ITC_STD" + "kur_amp_loff_v"+ "kur_amp_lon_g"	0.831	"skw_amp_loff_v" + "std_amp_loff_all"+ "kur_amp_loff_g"
DDW*[2] Volume (m^3)	-	-	0.550	"kur_ht_loff_v" + "amp95_loff_all"
No. Saplings	0.965	"amp60_lon_all"+ "ITC_ht" + "amp45_lon_all" + "var_amp_loff_v"	0.939	"AD_ht_lon_v" + "kur_ht_lon_all"
No. Seedlings	0.821	"DEM_slope_loff_all"+ "kur_amp_loff_g" + "var_amp_loff_v" + "amp15_lon_all"	0.460	"AD_ht_lon_v" + "kur_amp_loff_all"
Percentage bare earth	0.761	"amp20_lon_all"	0.850	"p05_lon_all" + "can_depth_loff_all" + "ITC_STD"

**[1]SDW = Standing Dead Wood; *[2]DDW = Downed dead wood; *[3]See Table 3 for explanation of codes.*

6. Discussion and conclusions

In agreement with numerous published studies (see Wulder *et al* 2012), this work has shown that the estimation of forest structural properties from airborne LiDAR data is feasible and may be used to derive estimations of forest vegetation structure over large areas after training with measurements from field plot information. When estimating plot-level metrics from LiDAR data it has been shown that significant correlations ($p < 0.05$) are possible using either DR or FW data.

However, even with the obvious shortcoming of a reduced pulse density, models produced from FW data overall produced results which accounted for more variance in the field data than models produced with DR data. Although the differences in explanatory power between the models produced for DR and FW were often small, they suggest that the acquisition of FW data may improve the accuracy of estimations of some forest variables (agreeing with the findings of Lindberg *et al* 2011) and enable the estimation of certain metrics which cannot be predicted with DR data alone.

A preliminary examination into the point clouds obtained over forested areas from the DR and FW datasets showed that the FW-derived point cloud had a greater overall number of returns per unit area, with a higher proportion in the mid-canopy and understorey. This agreed entirely with the findings of Chauve *et al* (2009) in black pine (*Pinus nigra*) forest, who generated 40-60% additional points from waveform decomposition, mainly located in the lower part of the canopy and in low sub-canopy vegetation. As with the current study, Chauve *et al* (2009) reported a negligible difference between FW and DR point cloud data for estimating canopy height, but an overall improvement in characterising the detailed vertical structure of forest stands in FW data.

Although validation with independent field data is still required, and taking into account the limited size of the training data (i.e. 21 field plots), these results suggest that full waveform LiDAR data describe the structural components of woodland to an equal (or better) level of accuracy than discrete return data, in spite of a lower spatial sampling rate due to the waveform digitisation process. As would be expected, the increased vertical profile information in full waveform LiDAR data enables a more accurate retrieval of within and sub-canopy structural components. Further study will examine the estimation of additional facets of vegetation structure, including the number of canopy layers and shrub layer characteristics, key information for assessing woodland condition and habitat quality.

Acknowledgements

We are grateful to the NERC Airborne Research & Survey Facility for the acquisition of the LiDAR data used in this study, and to Dr Peter Bunting (Aberystwyth University) and Dr John Armstron (University of Queensland) for access to and guidance with the use of their LiDAR processing software.

References

Anderson, J. E., Plourde, L. C., Martin, M. E., Braswell, B. H., Smith, M. L., Dubayah, R. O., Hofton, M. A., and Blair, J. B. 2008. Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment*, 112, 1856-1870.

Armston, J. RSC lidar tools. Available from: http://code.google.com/p/rsclastools/ [Accessed: 23/05/2012].

Bunting, P., Armston, J., Clewley, D., and Lucas, R., 2011a. The Sorted Pulse Software Library (SPDLib): open source tools for processing lidar data. In: *Proceedings of "SilviLaser 2011"*. Hobart, University of Tasmania: 615-625.

Bunting, P., Armston, J., Lucas, R., and Clewley, D., 2011b. Sorted Pulse Data (SPD) format: a new file structure for storing and processing lidar data. In: *Proceedings of "SilviLaser 2011"*. Hobart, University of Tasmania: 573-582.

Chauve, A., Vega, C., Durrieu, S., Bretar, F., Allouis, T., Deseilligny, M.P., and Puech, W., 2009. Advanced full-waveform lidar data echo detection: assessing quality of derived terrain and tree height models in an alpine coniferous forest. *International Journal of Remote Sensing*, 30, 5211-5228.

Chen, Q., Gong, P., Baldocchi, D., and Tian, Y. Q., 2007. Estimating basal area and stem volume for individual trees from lidar data. *Photogrammetric Engineering and Remote Sensing.* 73, 1355-1365.

Coops, N. C., Wulder, M. A., Culvenor, D. S., and St-Onge, B., 2004. Comparison of forest attributes extracted from fine spatial resolution multispectral and lidar data. *Canadian Journal of Remote Sensing*, 30, 855-866.

Falkowski, M. J., Evans, J. S., Martinuzzi, S., Gessler, P. E., and Hudak, A. T., 2009. Characterizing forest succession with lidar data: an evaluation for the inland northwest, USA. *Remote Sensing of Environment*, 113, 946-956.

Gamon, J. A., Field, C. B., Goulden, M. L., Griffin, K. L., Hartley, A. E., Joel, G., Penuelas, J., and Valentini, R., 1995. Relationships between NDVI, canopy structure, and photosynthesis in 3 californian vegetation types. *Ecological Applications*, 5 , 28-41.

Heinzel, J., and Koch, B,. 2011. Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, 13, 152-160.

Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., and Falkowski, M. J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from lidar data. *Remote Sensing of Environment*, 112, 2232-2245.

Hudak, A.T.. Evans, J.S., and Smith, A.M.S., 2009. LiDAR Utility for Natural Resource Managers. *Remote Sensing,* 1, 934-951.

Kalacska, M., Sanchez-Azofeifa, G. A., Calvo-Alvarado, J. C., Quesada, M., Rivard, B., and Janzen, D. H., 2004. Species composition, similarity and diversity in three successional stages of a seasonally dry tropical forest. *Forest Ecology and Management*, 200, 227-247.

Koukoulas, S., and Blackburn, G. A., 2004. Quantifying the spatial properties of forest canopy gaps using lidar imagery and GIS. *International Journal of Remote Sensing*, 25, 3049-3071.

Kronseder, K, Ballhorn, U., Bohm, V., and Siegert, F., 2012. Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data. *International Journal of Applied Earth Observation and Geoinformation*, 18, 37-48.

Lin, Y.C., and Mills, J.P., 2010. Factors Influencing Pulse Width of Small Footprint, Full Waveform Airborne Laser Scanning Data. *Photogrammetric Engineering and Remote Sensing*, 76, 49-59.

Lindberg, E., Olofsson, K., Holmgren, J., and Olsson, H., 2012. Estimation of 3D vegetation structure from waveform and discrete return airborne laser scanning data. *Remote Sensing of Environment*, 118, 151-161.

Newton, A. C., Cantarello, E., Myers, G., Douglas, S., and Tejedor, N., 2010. The condition and dynamics of New Forest woodlands. In: (Eds.) A.C. Newton. *Biodiversity in the New Forest*. Newbury, Berkshire: Pisces Publications: 132-147.

Persson, Å., Söderman, U., Töpel, J., and Ahlberg, S., 2005. Visualization and analysis of full-waveform airborne laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 103–108.

Pirotti, F., 2011. Analysis of full-waveform LiDAR data for forestry applications: a review of investigations and methods. *iForest-Biogeosciences and Forestry*, 4, 100-106.

Reitberger, J., Schnorr, C., Krzystek, P., and Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform Lidar data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, 561-574.

Vaughn, N.R., Moskal, L.M., and Turnblom, E.C., 2011. Fourier transformation of waveform Lidar for species recognition. *Remote Sensing Letters*, 2, 347-356.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., and Studnicka, N. 2006. Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60, 100-112.

Wagner, W., Hollaus, M., Briese, C., and Ducic, V., 2008. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. *International Journal of Remote Sensing*, 29, 1433-1452.

Wulder, M.A., White, J.C., Nelson, R.F., Naesset, E., Orka, H.O., Coops, N.C., Hilker, T., Bater, C.W., Gobakken, T., 2012. Lidar sampling for large-area forest characterization: a review. *Remote Sensing of Environment*, 121, 196-209.

Zhang, K. Q., Chen, S. C., Whitman, D., Shyu, M. L., Yan, J. H., and Zhang, C. C., 2003. A progressive morphological filter for removing nonground measurements from airborne lidar data. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 872-882.

Validation of stand-wise forest data based on ALS

Ola Lindgren[1]

[1]OL Skogsinventering AB, Kopmangatan 44 A, SE-83133 Ostersund, Sweden,
ola.lindgren@olskog.se

Paper Number: ### SL2012-014

Abstract

The paper describes a method of using independent field data for quality control within the framework of commercial ALS projects. A stand level validation of large (10-20 ha in our case) stands is not without difficulties. It will not be economically justified to enumerate large stands for the purpose of quality control. A traditional sample plot layout will introduce a random error component of considerable size compared to the random component associated with the ALS data itself. The validation should estimate the random errors in ALS data, as well as the bias. Thus, one has to find ways to avoid, or at least estimate and compensate for, the random error component from the validation survey itself. A validation model based on the estimation of variance components is described. The model was applied to three commercial ALS projects in Sweden. The results show that the standard error of stand-wise volume estimates of an ALS data set would be overestimated by 50 % if the validation data were treated as true values. The paper ends with a discussion of possible alternatives to our present approach.

1. Introduction

In a management information system for forestry, data such as volume ha^{-1}, basal area ha^{-1} or mean height are usually aggregated at stand or compartment level. From a management perspective, it is important to know the quality of such data, since decisions about forestry operations are made at this level. Accurate data will lead to better decisions, which may justify the costs of data acquisition within an ALS (Airborne Laser Scanning) project. In commercial projects, a statement of expected accuracy on stand level is usually part of the business agreement. Accordingly, there is a need for methods of assessing the data quality (validation) at stand level.

The need for stand level validation is recognized by most researchers in the field, but there is a considerable variation between methods. The most brutal method is to cut down the validation stands and compare the amount of extracted wood with ALS estimates (Barth et al.2008). A full enumeration of standing trees within real stands is usually considered too expensive, unless the stands are very small. A concept of “micro-stands” within actual larger stands was used by Naesset (2004) and Holmgren (2004), using micro-stands of 0.4 - 0.5 ha. This approach is sufficient if the “target objects”, i.e. the real stands being managed, are only somewhat larger.

The average stand sizes within company owned forests in northern Sweden are typically in the range 10 – 20 ha. To validate ALS data for such stands, a sampling approach within stands is needed. A standard method is to lay out a square grid of circular plots over each validation stand (Naesset 2002, Holmgren and Jonsson 2004). Such an approach is straight-forward for estimating a possible bias in the ALS data. Bias control is essential if the stand data should be used for purposes such as long term timber supply modelling.

However, a traditional sample plot layout within validation stands will introduce a random error component of the same, or even greater, magnitude than the random component associated with the ALS data itself. It is important to assess the magnitude of the random errors at stand level, since it determines the usefulness of the data for stand level decisions. If we treat the validation data as if they are "true", we will overestimate the expected random errors due to ALS at stand level.

The next sections of this paper will describe how we have carried out the validation part in three commercial ALS projects, and how we have dealt with the issue of isolating the random errors due to ALS data only. Finally, alternative methods and possible improvements will be discussed.

2. Method

2.1 Project areas

The validations have been carried out in three locations in Sweden, the characteristics of the areas are summarized in Table 1.

Table 1: Description of the project areas of the validation projects

Project (year)	Size, ha	Latitude	Altitude	Age classes in project	Mean volume, m3 ha^{-1}	Main species
Storådalen (2005)	4 700	61.5	500	31-	80	Pine (spruce, birch)
Hassela (2006)	7 000	62	390	21-	200	Pine, spruce, contorta (birch)
NV Dalarna (2010)	140 000	61	470	31-	132	Pine (spruce, birch)

2.2 Laser data and training data

All projects are applications of the two-stage area-based method as proposed by Naesset (2002). The laser point densities were in the range 0.5-1.0 points/m^2. In the first stage, forest parameters were estimated for grid cells ranging from 14 x 14 m (Hassela) to 18 x 18 m (Storådalen). In the second stage, cell estimates were combined to calculate estimates on stand level. The estimated parameters were volume ha^{-1}, basal areal ha^{-1}, stems ha^{-1}, mean diameter and mean height. In the most recent project (NV Dalarna), spectral information from aerial photos was also used in the estimation process to provide species-wise data (Packalén and Maltamo 2006).

The training data used for making the laser-based predictions consisted of slightly clustered circular plots with sizes corresponding to the grid cells used for the first stage of estimation. The number of plots were 112 (Storådalen), 181 (Hassela) and 503 (NV Dalarna). Rough stratifications based on dominant species and/or age class were used. The field plots were measured and processed using routines from the Forest Managament Planning Package (FMPP) (Jonsson et al. 1993).

2.3 Validation data

The validation stands were selected by stratified random sampling, with PPS-selection within strata. In the Storådalen project, new stand boundaries had been established by photo interpretation. In the Hassela project, old boundaries were kept intact while new boundaries based on laser data were created by a semi-automatic procedure in the NV Dalarna project. In each selected stand, circular plots were laid out using a square lattice. Table 2 summarizes the design characteristics.

Table 2: Description of the sampling design of the validation surveys

Project (year)	Average stand size, ha	No of stands	No of plots per stand	Plot raduis, m
Storådalen (2005)	26	10	15	10
Hassela (2006)	19	15	15	10
NV Dalarna (2010)	40	25	15	8-10

The sampling effort per stand (no of plots, radius) is quite ambitious compared to standard management inventories. As for the training data, sample plots were measured and processed by the FMPP routines.

2.4 Validation model

The validation model should be able to separate the random error component from ALS data from the random errors introduced by the sampling within the validation stands. A model from Ståhl (1992) was used as a basis. This approach was also used by Holmgren and Jonsson (2004).

Let

$$L_i = a + \eta_i + \varepsilon_i \quad (1)$$

where L_i = ALS estimate for stand i
a = a constant
η_i = the true value
ε_i = a random error with expectation zero

and

$$V_i = \eta_i + \delta_i \quad (2)$$

where V_i = Validation estimate for stand i
δ_i = a random error with expectation zero

Focusing the differences, we have

$$D_i = L_i - V_i = a + \varepsilon_i - \delta_i \quad (3)$$

Since the random error terms both have expectation zero, the unknown mean difference (a) is estimated by D, the arithmetic mean of the differences D_i. The variance of D is

$$\mathrm{Var}(D) = \mathrm{Var}(a + \varepsilon_i - \delta_i) = \mathrm{Var}(\varepsilon) + \mathrm{Var}(\delta) \quad (4)$$

since ε and δ are considered independent. The variance of the random error due to ALS is

$$\text{Var}(\varepsilon) = \text{Var}(D) - \text{Var}(\delta) \quad (5)$$

Var (D) is estimated from the observed series of differences, while the variance component from the validation survey, Var (δ) is estimated by

$$\text{Var}(\delta) = (1/n)\ \Sigma\ SE_i^2 \quad (6)$$

where SE_i = the standard error for the validation estimate for stand i

Since the validation surveys were made by systematic sampling and not by SRS (simple Random Sampling), the standard errors SE_i cannot easily be estimated from the data itself. The approach chosen was to reduce standard errors calculated by the SRS wr formula by applying a correction factor of 0.7. This is a quite rough figure based on sampling simulation in mapped stands (Lindgren 1984).

3. Results

Results will only be shown for the key variable of interest, volume ha^{-1}. Figures 1-3 show the graphics of the relations between stand-wise volume estimates from ALS and the corresponding estimates from the validation surveys.

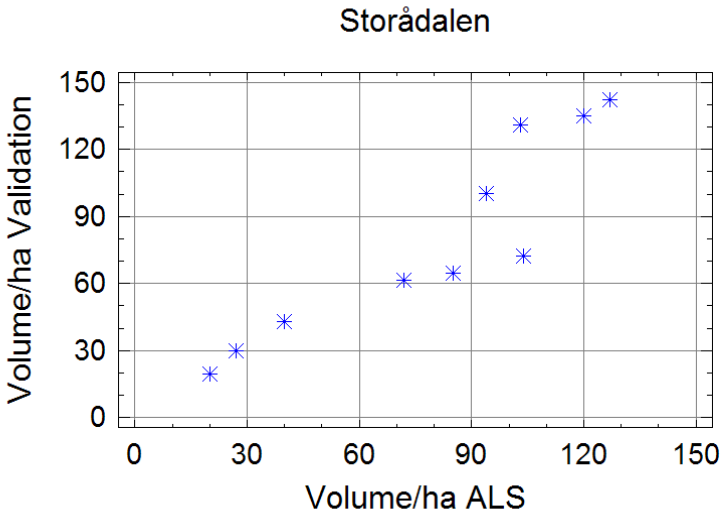


Figure 1: Comparative plot for the Storådalen project. Markers represent stands.

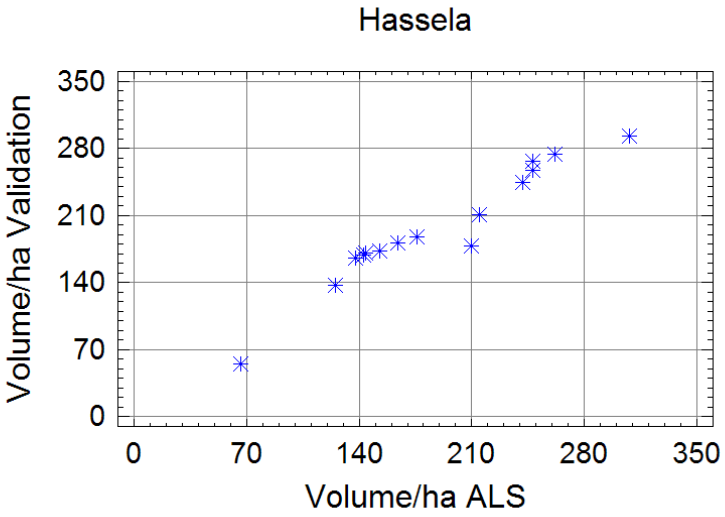


Figure 2: Comparative plot for the Hassela project. Markers represent stands.

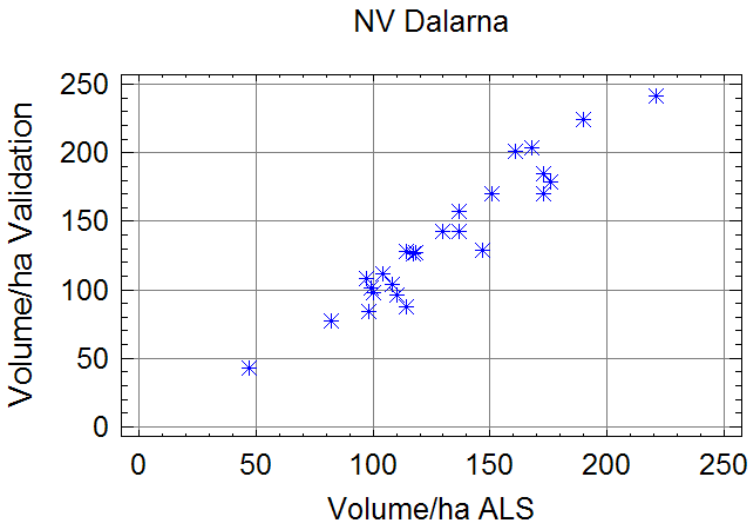


Figure 3: Comparative plot for the NV Dalarna project. Markers represent stands.

The linear correlations are quite high (r = 0.92, 0.96, 0.96). Note that random errors from the validation surveys are components of the variation as well as random errors due to ALS estimates. If the validation stands had been fully enumerated, still higher correlations would have appeared.

Table 3 shows the bias estimation of the volume validation.

Table 3: Evaluation of bias for volume ha^{-1}

Project (year)	Mean volume	Mean diffence (L-V)	Mean difference, %	St.Dev of differences, %	Prob-value for H_0 to be true
Storådalen (2005)	80	-0.71	-0.89	21.09	0.900
Hassela (2006)	200	-8.13	-4.07	8.56	0.087
NV Dalarna (2010)	132	-5.73	-4.34	12.46	0.095

The two latter projects show tendencies towards underestimation in the ALS data, although prob-values for H_0 (the null hypothesis) are not low enough to be certain of a real bias in the whole sampled population. Table 4 shows the total random errors, and the division of random errors into components.

Table 4: Random error components for volume ha^{-1} and their effect on the estimated standard errors of ALS data. Recall the validation model in section 2.4 for explanation of the variance components.

Project (year)	Var (D)	Var (δ), valid survey	Var (ε), ALS data	SE unadjusted	SE adjusted	Adjustment factor
Storådalen (2005)	283.92	161.29	122.63	16.85	11.07	0.66
Hassela (2006)	292.75	162.31	130.44	17.11	11.42	0.67
NV Dalarna (2010)	270.93	145.20	125.73	16.46	11.21	0.68

In spite of the quite ambitious validation (15 large plots per stand), the variance component from the validation survey was larger than the component from ALS data in all three projects. In terms of standard errors (SE), the relative effects of random errors from the validation survey decreases. If standard errors were calculated as if the validation data were "true", they would overestimate the adjusted standard errors with about 50 %.

Interestingly, the statistics for volume ha^{-1} are very similar for all three projects. Recall from table 3 that the average stocking varied considerably. If adjusted standard errors of ALS data are reported in % instead, we get 13.9 % for Storådalen, 5.7 % for Hassela and 8.5 % for NV Dalarna.

4. Discussion

Our present approach to validation of ALS data for large stands has been presented above. An important feature of the method is that it recognizes and corrects for random errors in the validation data itself. A weak link is the estimation of the standard errors for each validation stand, applying a rough correction factor for the effect of systematic sampling. A few alternatives and approaches to further development will now be discussed.

4.1 Using harvester data

Data from a modern well calibrated harvester will generate an accurate record of the tree removal after clear-felling of a stand. Not only will the total volume be given, but also the size distributions of the felled trees. The most obvious drawback of this method is its restricted application to old stands, and that the selection of stands is beyond control of the forest inventory specialist.

The method is not free from field work. The logging site has to be surveyed and the area calculated. The volume retained for nature conservation must be quantified. Furthermore, one has to verify that the actual stump heights and top diameters are in accordance with the volume estimation models used when estimating the ALS data. In spite of the restricted scope, it is not a bad idea to use harvester data in a follow-up system to get experiences of the relation between extracted wood and the pre-harvest data. However, deviations will not entirely be due to errors in pre-harvest data but be influenced by many other factors.

4.2 Using SRS of validation plots

Systematic sampling of plot locations within stands is usually applied when making objective stand inventories such as validation surveys. There are no generally applicable formulas to estimate the standard errors per stand in this case. Various re-sampling methods will not give reliable standard error estimates for small samples. Instead, we use experiences from simulation studies to try to correct for the increased efficiency of systematic sampling as compared to SRS (Simple Random Sampling). ALS data in the form of estimates by grid cells is an excellent data source for making such simulations.

To get an unbiased estimate of the variance component from the validation survey, one may consider using SRS instead of systematic sampling. This is easy to carry out using the present GIS/GPS capabilities. However, assuming a constant sample size, the gain in variance estimation will be followed by a less accurate bias estimation (detection) since systematic sampling generally gives more precise estimates than SRS.

4.3 Using ALS data for validation plots only

With our present approach, we compare ALS data for the whole stand with validation data from sample plots. In case of large stands, the area of sample plots will be a small fraction of the whole stand. Example: 15 plots with 10 m radius cover approximately 0.5 ha, while the average size of the stands in our validations is 28 ha. This is the source of the random error component from validation that we try to eliminate.

Instead, one may consider calculating ALS data for the chosen validation plots only. The validation will then be made by comparing exactly the same area estimated by ALS as well as by on-the-ground validation. This may be considered as a variant of the fully enumerated “micro-stand” concept (se section 1). The difference is that the enumerated 0.5 ha area is distributed over the whole stand instead of being concentrated to one large plot. Systematic sampling would be preferred since there will be no need to estimate variance components in this case. There are some drawbacks of this method, mainly concerning additional costs and increased time span needed to carry out the validation.

The validation plots must be accurately positioned, which is not the case with our present method. Accurate positioning will require post-processing in remote parts of Sweden since real-time enhancement services often are out of signal reach. Furthermore, an additional step of ALS calculations for the positioned validation plots is needed. If very accurate real-time positioning is available, this step might not be necessary. One could instead try to locate already processed grid cells – or aggregates of grid cells - in field and use them as validation plots.

Acknowledgements

Many thanks to the customers who have financed the projects: Bergvik Skog AB, Holmen Skog AB, Sveaskog AB and the Municipality Forests of Orsa, Lima, Transtrand and N Venjan. Our partners Prevista AS (Norway), Cowi AS (Norway) and Arbonaut OY (Finland) have processed the ALS data. Finally, many thanks to our field staff who has carried out the tough part of the work and encountered a lot of mosquitos and a few bears.

References

Barth, A., Hannrup B., Möller, J. and Wilhelmsson, L. 2008. Validering av FORAN SingleTree® Method. *Arbetsrapport nr 666, Skogforsk, Uppsala, Sweden. (In Swedish)*

Holmgren, J. 2004. Prediction of Tree Height, Basal Area and Stem Volume in Forest Stands Using Airborne Laser Scanning. *Scand. J. For. Res 19: 543-553, 2004.*

Holmgren, J and Jonsson, T. 2004. Large Scale Airborne Laser Scanning of Forest Resources in Sweden. In Laser-Scanners for Forest and Landscape Assessment, 03-06 October, 2004, Freiburg. Germany. ISPRS Archives *http://www.isprs.org/proceedings/XXXVI/8-W2/*

Jonsson, B., Jacobsson, J. and Kallur, H. 1993. The Forest Management Planning Package. Theory and application. *Swedish University of Agricultural Sciences, Studia Forestalia Suecica 189, Uppsala, Sweden.*

Lindgren, O., 1984. A study on circular plot sampling of Swedish forest compartments. *Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management, Report 11, Umeå, Sweden.*

Naesset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment 80 (2002) 88-99.*

Naesset, E., 2004. Practical Large-scale Forest Stand Inventory Using a Small-footprint Airborne Scanning Laser. *Scand. J. For. Res 19: 164-279, 2004.*

Packalén, P. and Maltamo, M.. 2006. Predicting the Plot Volume by Tree Species Using Airborne Laser Scanning and Aerial Photographs. *Forest Science 52(6) 2006: 611-621.*

Ståhl, G., 1992. A study on the quality of compartmentwise forest data acquired by subjective inventory methods. *Swedish University of Agricultural Sciences, Department of Biometry and Forest Management, Report 24, Umeå, Sweden. (In Swedish with summary in English)*

Toward an Enhanced Forest Inventory Using Terrestrial and Airborne LiDAR in Newfoundland: a Virtual Journey through the *L-Architect* Model

Jean-François Côté [1], Richard A. Fournier [2], Joan E. Luther[3] & Olivier R. van Lier[4]

[1]Natural Resources Canada, Canadian Forest Service - Canadian Wood Fibre Centre, Jean-Francois.Cote@NRCan-RNCan.gc.ca
[2]Centre d'applications et de recherche en télédétection (CARTEL), Département de géomatique appliquée, Université de Sherbrooke, Richard.Fournier@USherbrooke.ca
[3]Natural Resources Canada, Canadian Forest Service – Atlantic Forestry Centre, Joan.Luther@NRCan-RNCan.gc.ca
[4]Natural Resources Canada, Canadian Forest Service - Canadian Wood Fibre Centre, Olivier.VanLier @NRCan-RNCan.gc.ca

Paper Number: #SL2012-015

Abstract

LiDAR data and architecture models have potential to enhance forest inventories with fine-scale information on forest structure. At the individual tree scale, terrestrial laser scanners (TLS) provide detailed and accurate 3-D data. However, in dense forest canopies, TLS data is limited by the occlusion of the laser beam from opaque objects and movement generated by wind. At the stand level, airborne laser scanners (ALS) have proven effective for mapping forest structure, however, occlusion here also prevents a complete profiling of material distribution. Moreover, current commercial TLS and ALS systems operate with only one wavelength which limits the complete description of the canopy. In each case, methods to estimate structural attributes are generally based on empirical relationships between canopy attributes and metrics derived from the TLS and ALS data. In this study, we use an architectural model to overcome the limitations of current TLS and ALS systems in order to improve the fine-scale description of 3-D canopy structure. The architectural model *L-Architect* (LiDAR data to vegetation Architecture) allows fine-scale reconstruction of individual tree structure from TLS data. In this study, *L-Architect* was adapted and validated with structural measurements on branching structure and foliage distribution for two coniferous species found in Newfoundland (Canada), namely, balsam fir (*Abies balsamea*) and black spruce (*Picea mariana*). The model reproduced realistic tree structures using TLS data and allometric relationships to define the total amount of foliage, following two main steps: branch growth and the addition of foliage within the tree crown. *L-Architect*'s results compared favourably with *in situ* physical measurements of branching structure from the main stem and vertical distribution of foliage for 6 sampled trees. Mean correlations of 0.99, 0.94, 0.71 and 0.77 were obtained for branch diameter, branch length, angle of branch insertion, and vertical distributions of foliage respectively. A library of 233 individual trees was generated using the TLS-scans as input to the architectural model. *L-Architect* was applied to reconstruct 26 permanent sample plots in Newfoundland using the library of individual trees. Detailed 3D representations of forest canopies were derived and compared with TLS data for one plot and ALS data acquired over the 26 plots. A mean correlation of 0.93 and bias of 1.28 m were obtained between height percentiles of ALS first return and simulated (foliar and wood) material distributions. The detailed reconstruction of tree and plot structure provides important information that is relevant to forest ecological processes and may be linked to wood quality attributes.

1. Introduction

Recent economic challenges faced by the forest industry have resulted in an increased emphasis on the development of enhanced forest inventory tools for determining structure and quality attributes in order to maximize economic benefits from forests and ensure long-term sustainability of forest ecosystems. Remote sensing is an essential tool for mapping and monitoring the state and evolution of large forest ecosystems. Several studies propose methods for measuring forest attributes, but their applicability is often limited to small areas or specific scenarios because of the structural complexity of forest canopy. Thus, improving our mapping capability is tied to our knowledge of forest structure to identify its influences on remote-sensing images/observations.

Tree architecture describes the topological and geometrical development over time at the tree level (Barthelemy and Caraglio, 2007). The architecture characterizes at any given time the nature (type, shape) and relative arrangement (position, connection) of each of its components. The resulting three-dimensional (3-D) structure represents the expression of equilibrium between endogenous growth processes and exogenous constraints exerted by the environment. Therefore, information on fine tree structure can potentially be translated into key physiological attributes and also improve our understanding of tree functioning and habitats. The very high degree of variability among trees requires a large number of measurements to characterize adequately the 3-D structure at different scales. Consequently, only a small set of selected variables are measured in current forest inventories, e.g., crown dimension, diameter of the stem at breast height (DBH), and tree height (Avery and Buckart, 1983).

Terrestrial laser scanners (TLS) are tools adapted for measuring 3-D tree and forest structure. They provide detailed and accurate 3-D data rapidly and efficiently. An increasing number of methods are now proposed to link the TLS point clouds and structure: tree stem parameters (e.g. Maas *et al.*, 2008), individual tree LAI (e.g. Béland *et al.*, 2011), branching structure (e.g. Bucksch and Fleck, 2011) and to reconstruct plausible trees (Xu *et al.*, 2007). These applications rely on the capacity to provide accurate and reliable structure information which in turn depends on TLS data quality. Despite their accuracy, TLS data can rapidly become unreliable when used in natural forest environments. Occlusion of the laser signal can prevent complete measurement of a specific scene (van der Zande *et al.*, 2006). The potential of TLS to retrieve 3-D structure is also not fully exploited because of other problems, such as (i) movement generated by wind, (ii) presence of objects at a finer level than that which can be resolved by TLS, and (iii) misregistration of TLS position. Thus, our capacity to distinguish between wood and foliage material remains limited in these environments. Also, TLS scans provide a raw sketch of the spatial distribution of elements in 3-D but do not give specific information on their geometry and connectivity (topology).

The *L-Architect* (*L*iDAR to tree *Architect*ure) modeling tool was developed to address the limitations of TLS data in natural forest stands and to extract fine-scale structural attributes of individual trees (Côté *et al.* 2009, 2011). *L-Architect* was designed as a practical tool to synthesize and quantify spatial distribution of tree components from TLS scans. *L-Architect* uses geometrically registered TLS scans of individual trees to construct a model of the tree structure. Expanding the model to create a detailed representation of forest canopy structure requires the creation of a catalog of individual tree samples scanned with a TLS (Côté *et al.*, 2012). The results from *L-Architect* are well adapted to natural forests, but some improvements are required before it can become a fully operational tool. A current project in Newfoundland (Canada) aims at forest characterization and enhanced inventory incorporating wood-fibre attributes using LiDAR data and architectural modeling. These tools will allow precise estimation of key structural attributes not currently available in the inventory. *L-Architect* is used to provide detailed tree structure for calculating metrics related to tree and forest stand

structure. Although *L-Architect* has proven its capability in several natural forest environments, the main goal of this study was to enhance the model's capabilities. The first objective was to adapt *L-Architect* to reduce computing time required to generate plausible tree structure while improving some of the algorithms used in earlier versions. The second objective was to implement architectural models of balsam fir (*Abies balsamea* (L.) Mill.) and black spruce (*Picea mariana* (Mill.) B.S.P.) in *L-Architect* and to validate these models with destructive measurements of branching structure and foliage distribution. The final objective was to use *L-Architect* to construct a detailed 3-D structural model of permanent sample plots (PSP) dominated by balsam fir and black spruce in Newfoundland.

2. Methods

2.1 *L-Architect* architectural model

L-Architect was adapted from its previous version (see Côté *et al.*, 2009; 2011; 2012). Improvements to the original tool resulted in the removal of time-consuming pre-processing and model parameterization. The modeling tool is now capable of generating models of trees with only two input variables: DBH and total leaf area. *L-Architect* is designed to construct 3-D tree structure models using TLS scans. We took advantage of the *Open L-Systems* formalism (Měch and Prusinkiewicz, 1996) to control branch growth and foliage addition within the tree crown. Specific rules, called productions, for simulating tree development were implemented at the bud level (Fig. 1). In our model, the main trunk develops from the terminal bud (order 0). The lateral buds on an axis of order n may produce buds that become the terminal bud of axis at order $n + 1$. The orientation of internodes or shoots may be modified by two factors: optimal growth direction determined by environment and the tendency to maintain a preferred orientation with respect to gravity (tropism). Two processes were implemented within an *Open L-System* framework for providing useful information to construct models of realistic trees in natural forests: (i) determining **space availability** for *branch growth*, and (ii) estimating **light exposure** for supporting *foliage addition* with a resource allocation model.

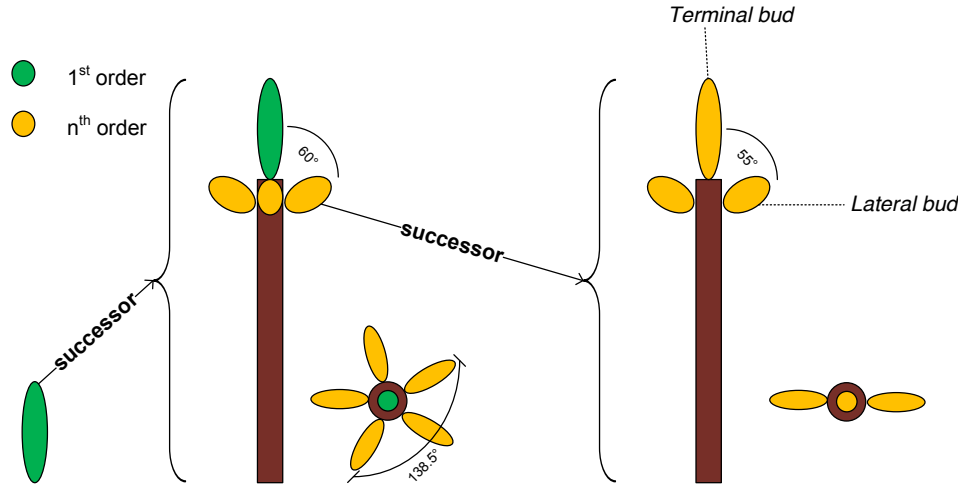


Figure 1: Schematic of bud formation. Two different successors (side and top view) are applied depending on if it is a terminal 1st order or a nth order bud.

The branching structure was developed using the productions protocols in conjunction with a space colonization algorithm (Runions *et al.*, 2007) implemented as an environment process. It assumes that each bud is surrounded by a spherical occupancy zone of radius ρ and has a

conical perception volume characterized by the perception angle θ and distance r (Fig. 2). The space available for tree growth is represented by a set of attractors S. These points are read from the TLS co-registered scans. At every step, attractors within the occupancy zone of any bud are deleted from the set S. An attractor within the perception volume of several buds is associated with the closest one. The optimal growth direction V of a bud A is calculated as the normalized sum of the normalized vectors toward all its attractors. This process is repeated until the attractor set is empty or it reaches a maximum number of steps. At the end, the branching structure of the tree is generated.

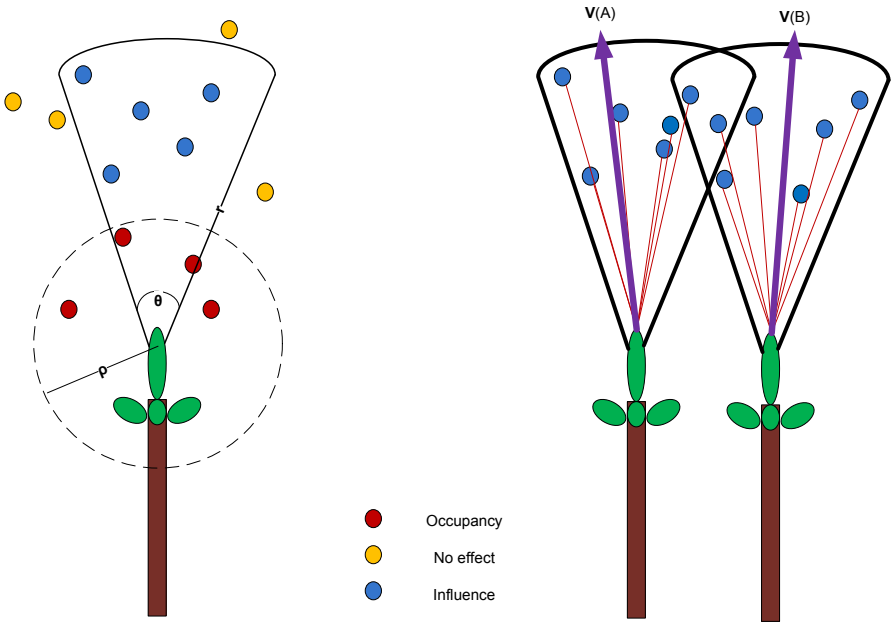


Figure 2: Space colonization. [Left] Perception volume and occupancy for a bud. [Right] Competition between remaining attractors. Adapted from Palubicki *et al.* (2009).

Foliage addition starts at the end of the branching structure development. It consists of adding and distributing a fixed number of conifer shoots in the tree crown. At each derivation step, we calculated the light availability at each bud position and used the cumulative light resource at the tree level to determine the buds' fate. A method based on shadow propagation (Palubicki *et al.*, 2009) was implemented to compute a coarse estimate of light exposure for each bud. Light availability served to determine which buds could produce new shoots. We implemented a priority resource allocation model that operates at the level of entire axes (Palubicki *et al.*, 2009). The total number of shoots to add into the tree structure was fixed at the beginning of the simulation and was proportional to total leaf area. Buds with the most available resources were chosen first (in descending order) to produce new shoots until we reached, in one or multiple derivation steps, the correct total leaf area.

We simulated branch shedding by comparing the total amount of resources gathered by a branch and the branch size measured in number of internodes. If both quantities fell below a specified threshold, the branch was considered a liability for the tree and was shed. Finally, branch diameters were estimated following a pipe modeling approach depending on the number of children a parent branch supported (Xu *et al.*, 2007).

2.2 PSP reconstruction with *L-Architect*

PSP reconstruction was accomplished by using the *L-Architect* model with supplemental measurements from inventory data. TLS scans were used to build a catalog of individual trees. We associated each scan with its species, DBH, total height (*H*), site class, and crown density class. For each tree found within the PSP, the algorithm chose a tree scan of the same species, site and crown density class from the catalog that minimized the following constraint between the scanned tree (*s*) and the measurement taken on the real tree in the inventory data (*i*):

$$\frac{|DBH_i - DBH_s|}{DBH_i} + \frac{|H_i - H_s|}{H_i}. \quad (1)$$

Once a tree from the catalog was chosen, we applied a scale factor to each point composing the TLS scan to obtain the same approximate height and crown dimensions of each tree found in the PSP. Then, DBH, *H* and total leaf area (*LA*) were given as inputs to *L-Architect*. The LA was calculated using allometric relationships developed specifically for balsam fir and black spruce trees in Newfoundland (Lavigne *et al.*, 1996; Newton, 1996). Foliage models were created using shoot and needle arrangements as measured for the selected species in a similar environment.

The procedures were repeated for all trees present on each PSP. Each tree of the PSP was positioned at the location identified from the inventory and was rotated randomly in azimuth to avoid spurious frequencies in the scene. Creating maps of tree locations required, first, a precise GPS location of the PSP central point. Then, the distance and azimuth of all trees in the PSP were measured from its center using a laser rangefinder. The (*x,y,z*)-positions of each tree were calculated using the GPS location, distance, and azimuth information.

2.3 LiDAR data

2.3.1 Terrestrial LiDAR

Measurements were conducted during summer 2009–2010 over 41 PSPs located in Newfoundland. The dominant species were balsam fir and black spruce. Georeferenced tree locations were measured within all the PSPs. A total of 209 individual trees located on different plots and one PSP comprising 57 balsam fir trees were scanned with a TLS: the Ilris-3D of Optech Inc. in 2009 and the Zoller+Fröhlich (Z+F) 5006h in 2010. The Ilris-3D emits light at 1,500 nm and scans with a maximum viewing window of 40°×40°, which can be adjusted to fit target objects. This system measures the position by calculating time delay between the emission and the reception of signal backscattering intensity in either first or last return mode. The Z+F emits light in the visible (680 nm) and scans with a near complete hemispheric viewing window of 310°×360°. It recorded return position using phase comparison.

At least two TLS scans were acquired from different positions to maximize viewing group of trees while reducing the negative impact of signal occlusion. The PSP scanned with the Z+F used nine different positions forming an "*X*" shape: one center location, four corner positions, and four between the center and corner positions. The 3-D point clouds taken by the TLS from different viewpoints were processed and aligned into one geometric coordinate system with the software Pointstream (3DImageSuite®).

2.3.2 Airborne LiDAR

A total of 1,193 km^2 of the commercial forest area of Newfoundland were flown in 2010–2011. Data were acquired with an Optech Airborne Laser Terrain Mapper (ALTM) 3100 system operating at a scan rate between 31–48 Hz with a scan angle of 20°. Acquisition was done at an altitude between 400–1,200 m with an approximate point density of 1–4 points/m^2. A ground classification was applied to the points in each block to clean the data of erroneous points and to compare the alignment of ground surfaces between flight lines. Airborne LiDAR data coincided with 26 PSPs reconstructed with *L-Architect*. Height-related metrics at 0.5-m resolution were extracted for each PSP and used to compare the vertical structure profiles derived from ALS data and simulated with the architectural model.

2.4 Tree level validation

We measured key attributes of branch geometry and foliage spatial distribution over a small but representative sample of trees to validate the model. A total of six trees, three balsam fir and three black spruce, were selected to address the validation exercise. For each species, selected trees represented the hierarchical class status of one regenerating, one immature (or juvenile), and one mature tree, all located in western Newfoundland (Canada). The balsam fir trees are hereafter referred to as bFregenerating, bFimmature, and bFmature. Similarly, the black spruce trees are referred to as bSregenerating, bSimmature, and bSmature. Sampling was performed in two steps. First, each tree was scanned with the Ilris-3D. The scans were taken at two opposite viewpoint locations around the tree using first return modes. Second, each tree was cut to measure a suite of structural attributes. The DBH and total tree height were measured. For each branch and twig connected to the main stem, we measured its (i) height of insertion from the main stem base, (ii) basal diameter, (iii) length, (iv) zenith angle of insertion in the stem, and (v) the number of shoots developed during the previous year of growth (1 year old). The comparative analysis with the model was performed by grouping classes of each evaluated variable. The branch diameter classes were grouped by intervals of 5 mm, 10 mm, and 15 mm for the regenerating, immature, and mature trees, respectively. Similarly, the branch length classes were grouped by intervals of 30 cm, 40 cm, and 50 cm, and the vertical distribution of shoots by intervals of 25 cm, 50 cm, and 100 cm. Branch insertion zenith angles were grouped at intervals of 30° for all trees.

2.5 Comparisons between LiDAR data and simulated structure

Destructive sampling was not an option for the evaluation of our reconstructed PSP structure, and no other independent baseline data set were available to validate our approach at the plot level. A preliminary assessment of the 3-D site reconstruction method with *L-Architect* for the PSP was done by comparing the simulated structure with LiDAR metrics: (i) the density index matrix derived from the co-registered point cloud of the PSP scanned with the TLS and (ii) the height percentile derived from ALS data for the 26 coinciding PSPs.

First, the co-registered point cloud of the PSP scanned with the TLS was *voxelized* and processed to assign to each voxel a *density index* Λ that is proportional to the surface density of the material. The method was developed first by Durrieu *et al.* (2008) and adapted to multiple TLS scans by Côté *et al.* (2011), and is called hereafter *LVox* (*L*idar to *Vox*els). Λ depends on the number of laser beam returns:

$$\Lambda = \frac{N_{inside}}{N_{theoretical} - N_{before}}, \tag{2}$$

where $N_{theoretical}$ is the theoretical number of laser beams entering a voxel, N_{before} is the number of intercepted laser beams before the targeted voxel, and N_{inside} is the number of laser beam returns in the voxel. We compared the vertical profile of Λ obtained with *LVox* and the simulated structure with *L-Architect*. The translation between the geometric and voxel representation of the simulated PSP was done by summing the total surface area of all of the geometric objects falling in each voxel. Both 3-D matrices were normalized to allow their comparison by dividing each voxel's value with the sum of all voxels' values in the 3-D matrix.

Second, the height percentiles were calculated from the ALS first returns and also from the simulated structure of the PSPs by using the vertical distributions of the material (foliar and wood) surface area.

3. Results

3.1 Validation of reconstructed trees with measurements of sample trees

Visual assessment of the sampled trees reconstructed with *L-Architect* compared with original photographs and the co-registered TLS point clouds showed good correspondence (Fig. 3). Tree DBH and height were within differences of zero and 30 cm respectively with measurements as they were given as input to the model. The distribution of branches per class of base diameter (Fig. 4), branch length, and zenith angle as well as the vertical profiles of the number of shoots (Fig. 5) compared well with *in situ* measurements. Branch basal diameter distributions showed strong correlation ($r > 0.97$) with field measurements with a normalized root mean square error (NRMSE) between 0.03 and 0.40 (Table 1). Similar results were found for branch length distribution ($r \geq 0.88$; $0.09 \leq \text{NRMSE} \leq 0.23$), except for the bFmature where the reconstructed structure exhibited longer branches on average compared with the actual trees, with a low r value (0.30) and large NRMSE (0.37). The angle of insertion was highly correlated ($r \geq 0.77$) with field measurements and $0.04 \leq \text{NRMSE} \leq 0.30$ except for the regenerating and immature balsam firs. The bFregenerating showed a median value around 120° compared with 90° of the actual tree. A potential source of error in the insertion zenith angle measurement includes the improper z-alignment of the tree during the measurements of branch insertion angle. In the case of the bFimmature, 40% of the total number of branches, mostly small with a basal diameter of 2 mm, were missing and could explain in part the lack of branches oriented around 90°. Lastly, the vertical distribution of shoots between sampled and simulated trees was highly correlated ($r \geq 0.71$) for most of the trees; NRMSE ranging from 0.13 to 0.22. Lower r values were found for bSregenerating (0.512) and bFmature (0.683). Vertical distribution was shifted up for the reconstructed bSregenerating, whereas a shift down was observed for the bFmature. Nonetheless, both gave a NRMSE below 0.30.

Figure 3: Example of reconstructed tree with *L-Architect*. The mature balsam fir: [left] the original photograph, [center] the co-registered TLS point cloud, and [right] the reconstructed tree.

Table 1: Correlation between selected variables from field measured and reconstructed trees.

	Branch Diameter Distribution		Branch Length Distribution	
	r	NRMSE	*r*	NRMSE
bSregenerating	1.000	0.025	0.998	0.094
bSimmature	0.994	0.203	0.952	0.185
bSmature	0.978	0.204	0.918	0.146
bFregenerating	0.979	0.404	0.960	0.196
bFimmature	0.992	0.230	0.884	0.231
bFmature	1.000	0.171	0.303	0.369
	Branch Angle Distribution		Vertical Distribution of Shoots	
	r	NRMSE	*r*	NRMSE
bSregenerating	0.995	0.038	0.512	0.310
bSimmature	0.999	0.116	0.709	0.224
bSmature	0.769	0.268	0.946	0.131
bFregenerating	0.146	0.468	0.921	0.163
bFimmature	0.553	0.321	0.884	0.157
bFmature	0.776	0.299	0.683	0.268

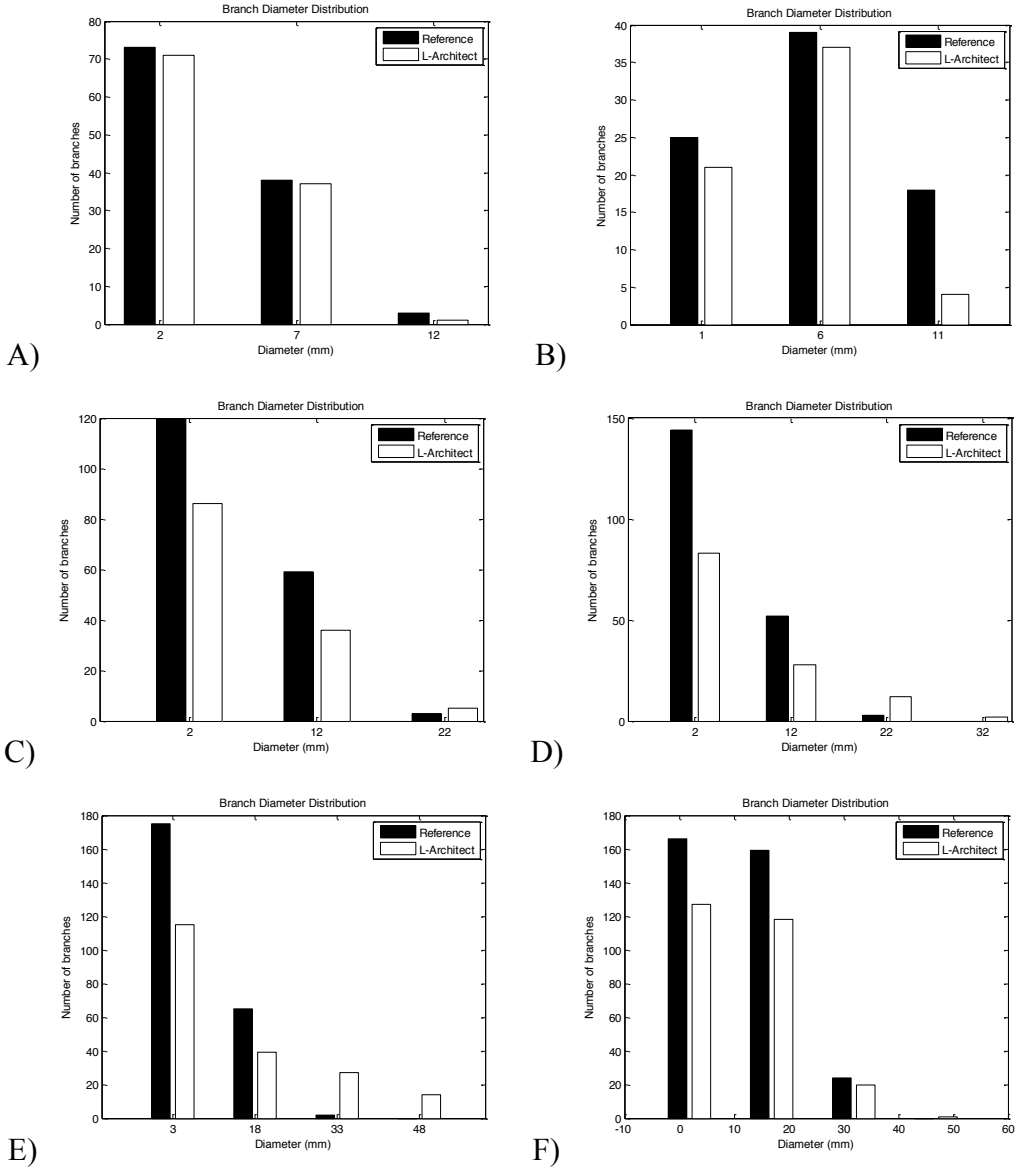


Figure 4: Branch diameter distribution for the reference field measurements (black) and reconstructed tree with *L-Architect* (white). A) bSregenerating, B) bFregenerating, C) bSimmature, D) bFimmature, E) bSmature and F) bFmature.

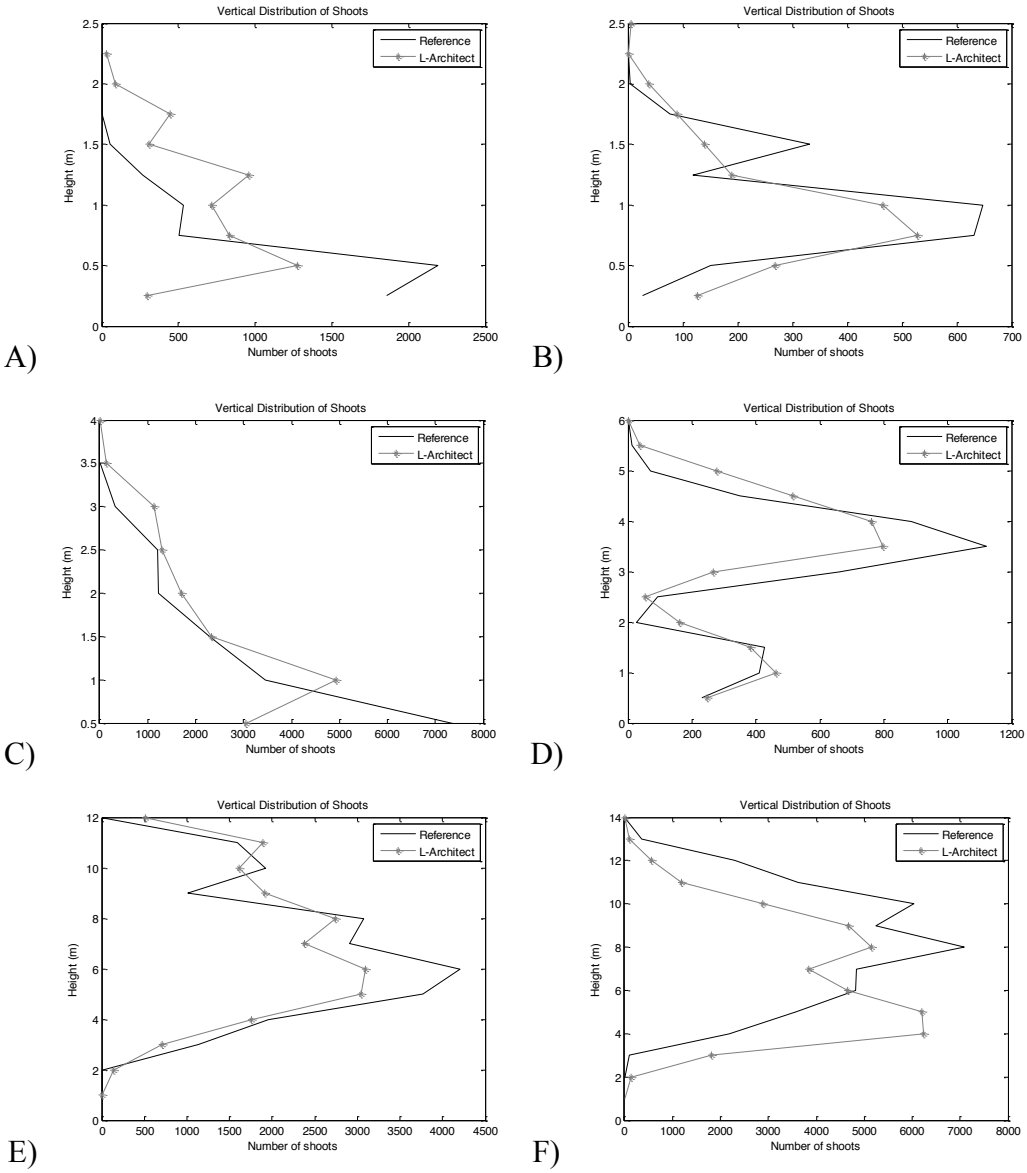


Figure 5: Vertical distribution of number of shoots for the reference field measurements (black) and reconstructed tree with *L-Architect* (grey - asterisks). A) bSregenerating, B) bFregenerating, C) bSimmature, D) bFimmature, E) bSmature and F) bFmature.

3.2 Evaluation of reconstructed PSPs with TLS and ALS

The comparison between the vertical normalized distribution of the simulated structure with *L-Architect* (Fig. 6) and TLS-derived density index matrix calculated with *LVox* for the scanned PSP is shown in Fig. 7. Correlation between the density matrix and the simulated structure for the vertical distribution was 0.86 with NRMSD = 0.0012. The Taylor diagram (Taylor, 2001) in Fig. 8 summarizes graphically how closely the simulated distribution of height percentiles matches observations, quantified in terms of their statistical correlation, centered root mean square deviation (CRMSD), and magnitude of the variability estimated by the standard deviation. Strong correlations were observed for all 26 PSPs between the distribution of height percentiles calculated from the ALS first returns and the simulated vertical profile of surface areas. A mean correlation of 0.9253 was observed where the bias never surpasses 3.21 m (mean

of 1.28 m). The differences can be related to (i) object occlusion inherent in a LiDAR data set, (ii) the fact that understory vegetation and fallen trees were not considered in the simulations, and (iii) positioning error of the simulated trees.

Figure 6: Co-registered point cloud (left) and rendered image (right) of the reconstructed PSP with *L-Architect* architectural model.

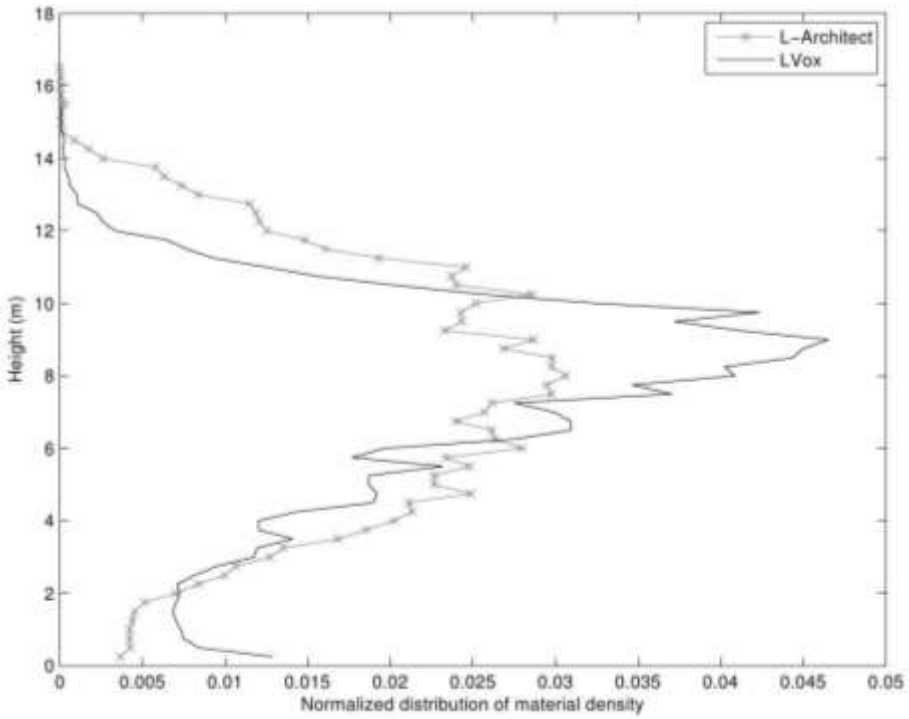


Figure 7: Vertical normalized distribution of material density for the *LVox* results and *L-Architect*.

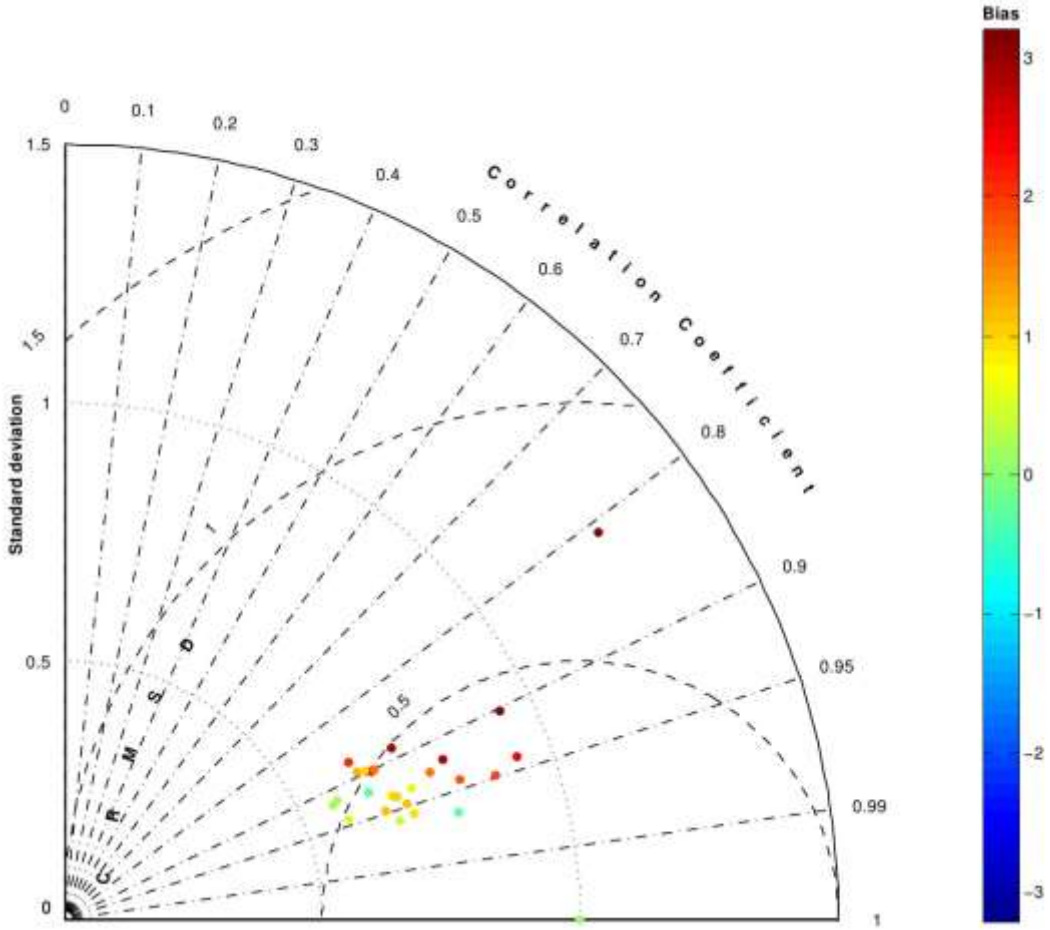


Figure 8: Taylor diagram comparing height percentiles from ALS data and *L-Architect* for 26 PSPs.

4. Discussion

The *L-Architect* model as adapted in this study produced realistic tree structure using TLS data and allometric relationships to define the total amount of foliage in two main phases: branch growth and foliage addition within the tree crown. *L-Architect* was successfully improved to recreate realistic individual balsam fir and black spruce trees within the constraint of practical application: the new version of the algorithm took a few minutes, compared with hours using the original model (Côté *et al.*, 2011). The comparisons with actual trees from different hierarchical status showed the capacity of *L-Architect* to model these new implemented species in boreal environments. The validation showed that *L-Architect* can characterize the detailed 3-D distribution of wood and foliage, which is generally impractical to do solely from analysis of TLS data in the context of natural mature forest environments. *L-Architect*'s results compared favorably with *in situ* measured branching structure from the main stem (diameter, length, angle of insertion) and vertical distribution of foliage. It is still unclear how better estimate on quality attributes can be achieved through the use of fine scale structure information (Fernández *et al.*, 2011). However, crown shape, branching structure and foliage distribution are very likely to be key factors in predicting the resulting attributes. The model expansion toward a detailed representation of forest canopy structure required the creation of a catalog of 209 individual tree TLS scans. Though preliminary, the analysis between LiDAR metrics and *L-Architect* structure has shown the potential of the modeling approach to replicate the complexity of mature natural conifer forest canopies. The creation of surrogate ground-reference will be useful for reducing the cost of ground measurements without compromising the quality of the attribute estimates.

5. Conclusions

L-Architect is an effective way to remove current limitations of TLS data sets affecting the fine-scale characterization of tree and plot structure, such as object occlusion, wind, and wood–foliage separation. Our approach reduces substantially the need for destructive measurements of tree structure while providing reliable estimates of forest structural characteristics, beyond what can be realistically done manually in the context of forest inventory. *L-Architect* could thus support in the future the development of relationships between tree structural attributes and other biophysical aspects such as growth or wood fiber quality. The ability to use these structural attributes is especially important to enhance current and future forest inventories. Improved implementation of *L-Architect* also permits the use of the model operationally and to reconstruct complete 3-D canopies with TLS scans to augment the structural information extracted from airborne LiDAR data and high-resolution optical imagery.

Acknowledgments

We would like to thank Danny Blanchette, Janet Bourque, Matthew Glover, Heidi Kavanagh, Emilie Lessard, and Michael Noonan for their help acquiring and processing TLS scans and measuring tree structure. We are very grateful to Tim Moulton of Corner Brook Pulp and Paper Limited for providing site access and permission to cut sample trees on company land.

References

Avery, T.E., and Burkhart, H.E. 1983. *Forest measurements*. McGraw-Hill series in forest resources. McGraw-Hill Book Company, New York.

Barthélémy, D., and Caraglio, Y. 2007. Plant architecture: a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. *Annals of Botany*, 99, 375–407.

Béland, M., Widlowski, J.-L., Fournier, R.A., Côté, J.-F., and Verstraete, M.M., 2011. Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements. *Agricultural and Forest Meteorology*, 151, 1252–1266.

Bucksch, A., and Fleck, S. (2011). Automated detection of branch dimensions in woody skeletons of fruit tree canopies. *Photogrametric Engeneering and Remote Sensing*, 77, 229–240.

Côté, J.-F., Widlowski, J., Fournier, R.A., and Verstraete, M.M. 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial LiDAR. *Remote Sensing of Environment*, 113, 1067–1081.

Côté, J.-F., Fournier, R.A., and Egli, R. 2011. An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. *Environmental Modelling & Software*, 26, 761–777.

Côté, J.-F., Fournier, R.A., Frazer, G., and Niemann, O.K. 2012. A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure. *Agricultural and Forest Meteorology*, 166–167, 72–85.

Durrieu, S., Allouis, T., Fournier, R.A., Véga, C., and Albrech, L., 2008. Spatial quantification of vegetation density from terrestrial laser scanner data for characterization of 3d forest structure at plot level. In *SilviLaser 2008*, Edinburgh, UK, 17-19 September.

Fernández M.P., Norero A., Vera J.R., and Pérez E., 2011. A functional–structural model for radiata pine (Pinus radiata) focusing on tree architecture and wood quality. *Annals of Botany*. 108, 1155–1178.

Lavigne, M.B., Luther, J.E., Franklin, S.E., and Hunt, E.R. 1996. Comparing branch biomass prediction equations for *Abies balsamea*. *Canadian Journal of Forest Research* 26, 611–616.

Maas H.G., Bienert A., Scheller S., and Keane E., 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. *International Journal of Remote Sensing*, 29, 1579–1593

Měch, R., and Prusinkiewicz, P., 1996. Visual Models of Plants Interacting with Their Environment. In: *Proceedings of SIGGRAPH 1996*, 397–410.

Newton, P.F. 2006. Intraspecific competition processes and their management within black spruce (*Picea mariana* (Mill) B.S.P.) stands. Ph.D. dissertation, Faculty of Forestry, University of British Columbia, Vancouver, B.C.

Palubicki, W., Horel, K., Longay, S., Runions, A., Lane, B., Měch, R., and Prusinkiewicz, P. 2009. Self-organizing tree models for image synthesis. *ACM Transactions on Graphics*, 28, 1–10.

Runions, A., Lane, B., and Prusinkiewicz, P., 2007. Modeling Trees with Space Colonization Algorithm. In: *Proceedings of the Eurographics Workshop on Natural Phenomena 2007*, 63–70.

Taylor, K.E., 2001. Summarizing multiple aspects of model performance in single diagram. *Journal of Geophysical Research*, 106, 7183–7192.

van der Zande D., Hoet W., Jonckheere I., van Aardt J., and Coppin P. 2006. Influence of measurement set-up of ground-based LiDAR for derivation of tree structure. *Agricultural Forest and Meteorology*, 141, 147–160.

Xu, H., Gossett, N., and Chen, B., 2007. Knowledge and Heuristic-Based Modeling of Laser-Scanned Trees. *ACM Transactions on Graphics,* 26.doi: 10.1145/1289603.1289610.

LiDAR Urban vegetation transmission estimates and implications for energy policy

Thoreau Rory Tooke & Nicholas C. Coops

[1]Forest Resources Management, University of British Columbia, 2424 Main Mall, Vancouver BC, V6T 1Z4 rorytooke@gmail.com
[1]Forest Resources Management, University of British Columbia, 2424 Main Mall, Vancouver BC, V6T 1Z4 nicholas.coops@ubc.ca

Paper Number: ###SL2012-021

Abstract:

Trees play a contrasting role in regulating a building's energy balance. Trees can provide shade during summer months, reducing the need for space conditioning, while alternatively, shaded rooftops ultimately restrict the potential to generate energy from solar-based technologies. Understanding this interaction of tree shade on domestic energy supply and consumption has important implications when designing local energy policy. In this study we develop an approach to help guide the design of building and landscape specific energy policy using airborne LiDAR (light detection and ranging). Our study is divided into 2 parts. The first part discusses the option of integrating a shape-based transmission profile to account for the semi-transparent nature of vegetation in irradiance models. The second part then uses this technique to assess the effects and sources of tree shade on building walls and rooftops. Results indicate that most of the solar radiation attenuation from trees occurs on public space and neighbouring lots. The implications and applications of these outcomes are discussed in terms of energy policy, specifically solar access and building energy standards.

1. Introduction

The increasing availability of light detection and ranging (LiDAR) datasets to municipalities has encouraged new mapping initiatives aimed at informing the available shortwave irradiance within urban environments. Urban climatologists have long been interested in irradiance as a key driver of the urban heat island effect (Oke, 1982). Furthermore, emerging energy policies designed to promote the adoption of household renewable energy technologies have led to an increased interest in detailed spatial modeling of urban irradiance. As a result, the spatial prediction of photovoltaic and solar thermal suitability dominates current municipal LiDAR applications related to energy planning.

Nonetheless, it is important to acknowledge the contrasting role trees play when considering energy management in cities. While an automated assessment of the absence of shading on building rooftops across a city is the primary objective of most solar mapping applications (e.g. http://www.geoweb.dnv.org/applications/solarapp/, http://gis.cityofboston.gov/SolarBoston/, http://sfenergymap.org/, http://nycsolarmap.com/), tree shade on a building envelope, especially its walls and windows, has also been promoted as a low-cost option for reducing the energy consumed for space conditioning. Results of simulations by Simpson and McPherson (2006) indicate that the strategic planting of several trees can produce up to a 50% decrease in cooling energy use and a 23% decrease in peak electricity use. Similarly, Huang et al. (1987) show that a 25% increase in tree cover results in a 25 – 40% decrease in building cooling energy use, and Laband and Sophocleus (2009) demonstrate that a house in full sun uses 2.6 times the cooling energy as an identical building in dense shade. This potential

to reduce electricity consumption also encourages the added secondary benefit of using shade trees to mitigate carbon emissions. Akbari (2002), for example, explains that a shade tree can offset the carbon emissions associated with power plants at a rate of up to 4 times its physiological ability to sequester carbon.

Given the contrasting effects of trees on energy conservation and solar-energy supply, effective policy requires thoughtful design to balance these competing factors. While vegetation management strategies are already often integrated with municipal planning initiatives, two energy-specific policies currently exist that help to illustrate the juxtaposition of tree shade in household energy management strategies, namely solar access laws and building energy standards. Solar access, which is often closely related to solar easement laws, expresses the right of a building owner to receive sunlight across property lines without obstruction from another's property. Solar access may be used for various purposes, although the primary objective is typically to protect a building owner's investment in a solar-energy technology (photovoltaic or solar thermal), which is maximized with greater amounts of received sunlight. Some form of solar access, including solar easement, is currently implement at a state or local level in most US states, while California is often cited for having the most comprehensive solar access policy, known as the Solar Shade Control Act (sections 25980-25986 of the California Public Resources Code).

Typically in contrast to solar access policies, building energy codes and standards (including energy performance certificates), implemented at various jurisdictional levels, require or suggest tree-planting strategies to reduce the cooling energy requirements of a building. Although it may be possible to design a tree shading strategy that can satisfy both an energy conservation initiative while maximizing sunlight on local and neighbouring rooftops, such a situation requires a detailed analysis of the local environment and form. In fact, most simulation studies of tree shade for energy conservation as well as many of the municipal solar mapping initiatives omit or neglect the detailed spatial representation of urban form, which can have potentially negative consequences on the intended design outcomes of various energy policies.

At the scale of an entire city, irradiance models that integrate a spatially contiguous and detailed representation of the surface heights provide current best estimates of incoming sunlight and shade. Nonetheless, these geospatial irradiance models also often simplify all surface features as opaque objects, producing a 2.5 dimensional digital surface model that is then used to compute slope, azimuth and shading obstructions (Rich, et al., 1994; Kumar et al., 1997). Although buildings can be generalized as solid objects, vegetation presents a semi-transparent feature ideally suited to a more dynamic assessment of radiation transmission in relation to the local solar coordinates and seasonal changes in vegetation structure. Since the LiDAR returns act as a direct measure of light attenuation through the foliage it provides an optimal tool for advancing irradiance models to account for vegetation transmission. Furthermore, analysis of LiDAR point clouds in forests demonstrate interactions between the radiative properties and point-based representations of the vegetation canopy (Lovell et al., 2003; Riaño et al., 2003; Riaño et al., 2004; Coops et al., 2007), encouraging significant potential for application to urban shading assessments.

The purpose of this paper therefore presents two primary and related objectives. The first objective is to demonstrate the use of a shape-based vegetation transmission estimate derived from airborne LiDAR and its application to irradiance modeling in an urban setting. In this first part of the paper, a modified gap-probability approach is presented to relate the vertical distribution of vegetation structure with radiation transmission and then compared to LiDAR height and cover estimates for deciduous and evergreen tree species under leaf-on and leaf-off conditions. The second objective is to integrate the vegetation transmission results with an assessment of tree shading on building roofs and walls. The second part of the paper will also address the source of radiation attenuation, comparing public, local and private

neighbouring spaces. Discussion is focused on the application of spatially detailed modeling efforts to help better inform and design local energy policy related to tree shade.

2. Methods

2.1 Study Area and Data

The focus for this study is a 400 m by 400 m area within a treed residential neighbourhood in Vancouver, Canada. This residential area includes a mix of tree species (both deciduous and evergreen) at various heights up to 25 m. Airborne LiDAR data was acquired at two time periods: the first acquisition in early March 2007 under leaf-off conditions using a TRSI Mark II discrete-return sensor (pulse rate: 50 kHz, platform altitude: 800 m, wavelength: 1064 nm, and beam divergence angle of 0.5 mrad); and the second in late August 2011 under full leaf-on conditions using a Leica ALS60 sensor (pulse rate: 113.7 kHz, platform altitude: 750m, wavelength: 1064 nm, and beam divergence angle of 0.22 mrad). For both datasets ground and non-ground LiDAR returns were pre-classified by the provider. All trees within the study area were manually delineated and classified as either evergreen or deciduous using aerial imagery in addition to a 1 m gridded layer of the first return LiDAR hits. Only those cells identified as trees in both the leaf-off and leaf-on datasets were selected for the analysis (Figure 1). Additional spatial datasets include lot parcels and building footprints, both provided by the City of Vancouver.

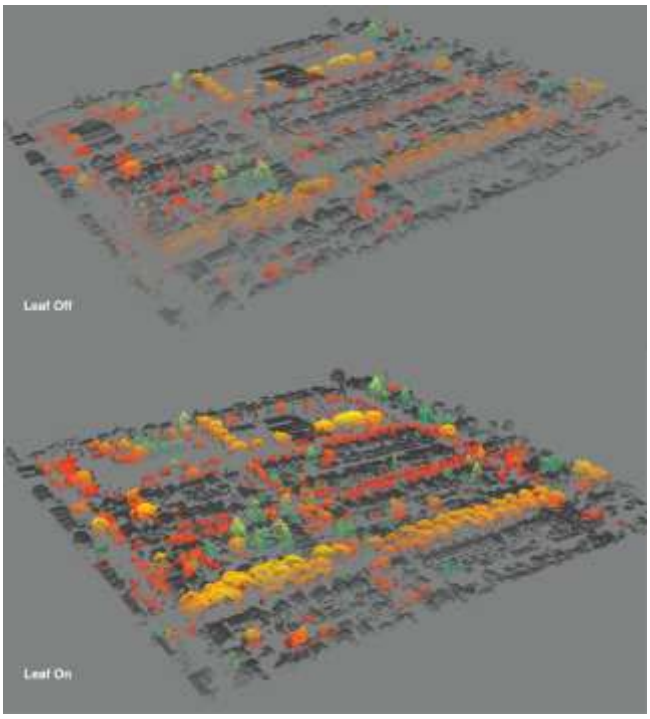


Figure 1. Three-dimensional representation LiDAR returns for leaf-off and leaf-on datasets within the study area with identified deciduous (reds) and evergreen (greens) tree types.

2.2 Shape-Based Transmission Profile

In forestry applications LiDAR data point clouds have been used to determine vegetation structural information (Lovell et al., 2003; Riaño et al., 2003; Riaño et al., 2004; Coops et al., 2007) and metrics have been developed based on three overall criteria: 1) height based metrics such as the coefficient of

variation for LiDAR returns, 2) metrics of the fraction of ground returns (which provides an indication of cover) and 3) shape based metrics which model the overall vertical profile of the crown. In the following section, all three criteria are compared for both LiDAR datasets capturing the same area. While both height and cover based metrics are well established, less research has taken place on curve fitting based approaches especially in urban environments.

Shape-based metrics of vegetation structure are of particular interest when examining radiation transmittance since they enable assessment of the dynamic interaction between radiation, tree structure and changing sun angles. One well-developed approach to transmission curve-fitting involves the derivation of gap probability models, which determine canopy structure by correlating the attenuation of the returned laser pulses with the density, size, and distribution of foliage and woody elements (Ni-Meister et al., 2001), producing a cumulative vertically projected profile of the foliage for an entire canopy.

The approach described in this study uses a modified version of this gap probability approach, with demonstrated good fit for various trees in an urban setting (*RMSE < 0.05*) (Tooke et al., 2012). A brief summary of the technique is provided below. The first step calculates a vegetation extinction coefficient (V_{ec}) as the inverse of the proportion of LiDAR returns at defined height intervals by:

$$V_{ec} = 1 - \frac{\sum_{i,j} \#z_{K_{v_{i,j}}}}{N_{K_v}} \quad (1)$$

where K_v is a subset of cells previously identified as vegetation, $\#z$ is the number of LiDAR returns at a height z above the ground, and N is the total number of above-ground LiDAR hits. To characterize vegetation structure for a wide range of tree species the second step then fits a two-parameter Weibull probability density function (W_{pdf}) to the extinction profile as:

$$W_{pdf} = \frac{\beta}{\alpha} \cdot \left(\frac{z}{\alpha}\right)^{\beta-1} \cdot \exp\left(-\frac{z}{a}\right)^{\beta} \quad (2)$$

where α is a vertical scaling parameter and β is a shape parameter that alters the breadth of the distribution. For each cell in the 1 m gridded vegetation classification layer the transmissivity profile (W_{pdf}) is calculated based on the non-ground LiDAR returns within a 3 by 3 cell neighbourhood window. As a result, α and β parameters are derived for each cell, which are intended to integrate with established geospatial irradiance models to account for vegetation transmission with varying local solar coordinates.

2.3 Integration with Spatial Irradiance Models

Spatially-explicit models of irradiance have been developed over several decades (Gueymard & Myers, 2008), with more recent applications to urban environments emerging alongside advances to contiguous spatial representations of surface heights (such as LiDAR). The various processes involved in solar modeling have been well described in the literature (Gueymard & Meyers, 2008; Hofierka & Suri, 2002), and can be separated into three basic components: solar geometry, atmospheric effects, and surface effects. Of these three components, surface effects are of particular interest in an urban environment due to its heterogeneous three-dimensional (3D) form.

Surface effects traditionally include obstructions from nearby features that occlude solar angles above the horizon. However, additional surface effects can also include the reflectance as well as the partial attenuation of radiation through semi-transparent features such as vegetation. In the case of basic obstructions, irradiance models apply a hemispherical viewshed algorithm (Rich et al., 1994), in which the maximum angle of obstructing features are assessed from a digital surface model for a set of radially

projected vectors between 0 and 360°. As the sun moves across the sky the local solar coordinates are then compared to the nearest obstruction to determine whether the surface is shaded. For computational efficiency and integration with input surface models, spatial approaches typically use an array-based structure and compute irradiance for each cell. It is within this modeling context that the integration of a gap probability transmission profile is discussed.

Using the basic approach of projecting vectors between the surface cell and nearby obstructions, the transmission profile of trees is integrated into the model using the α and β parameters discussed in section 2.2. In this case trees must be differentiated from solid features in the input digital surface model. Then for a given local solar coordinate (solar azimuth and elevation), the extinction coefficient is computed from equation 2 using α and β, and z representing the height of the sun at the intersecting cell. At each solar coordinate during the model iterations the total extinction is computed as the product of all tree cell extinction values at those angles greater than the solid obstructions.

2.4 Place-Based Assessment of Vegetation Attenuation

As an example of the application of the above approach to an analysis of energy policy, we compute the radiation attenuation through the tree canopy of the study area over an entire year. The model runs are executed at hourly intervals for the 15th day of each month and separated into attenuation on building walls and building roofs. Using parcel data for the City of Vancouver, the vegetation attenuation associated with both roofs and walls is then further assessed to determine the source of the attenuation. Attenuation sources are divided into three basic place types, each with important policy design implications. Attenuation on the lot on which a given building is situated is the first type, referred to hereafter as *local*. Next, *private* refers to the attenuation associated with private parcels excluding those trees on the building lot itself. Finally, *public* refers to the radiation attenuation from trees on public space, notably street trees.

3. Results

The manual classification of tree types within the study area identified the total plan area of deciduous trees ($\lambda_{deciduous,T}$) at 17.4% and evergreen trees ($\lambda_{evergreen,T}$) at 11.7% of the study area. Figure 2 shows the difference in the coefficient of variation of LiDAR returns and fraction of ground returns. The mean coefficient of variation for both leaf-on and leaf-off datasets is approximately equal irrespective of tree type, although the variance is slightly reduced in the leaf-on dataset. The fraction of ground-returns provides a surrogate for cover, and as expected, is higher in the leaf-off dataset (approximately 35%) than the leaf-on dataset (approximately 25%) with cover slightly higher for evergreen tree types.

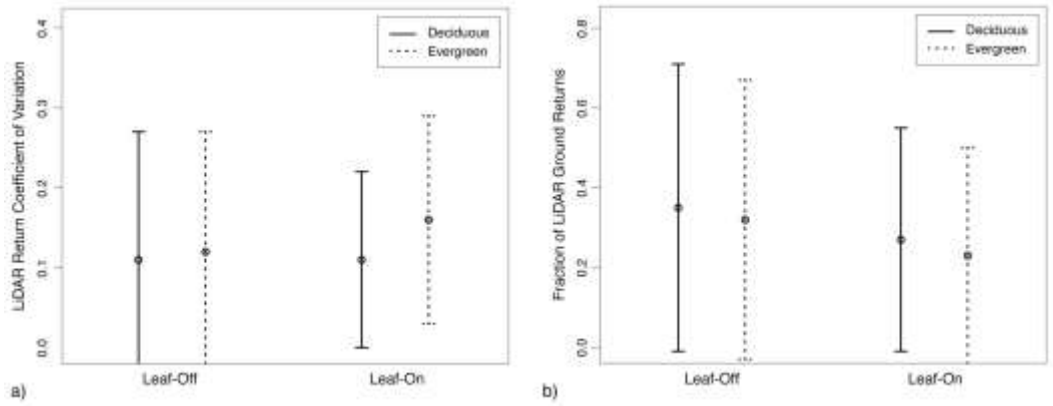


Figure 2. Comparison of mean and standard deviation for a) coefficient of variation and b) ground returns between the leaf-off and leaf-on LiDAR datasets.

The results of the shape-based structural assessment of transmission are shown in Table 1 and Figure 3 and highlight the differences in the vertical distribution of vegetation structure between datasets and tree types. In leaf-off conditions deciduous trees produce a vertical profile with a much narrower breadth compared to leaf-on conditions, while the shape of the evergreen distribution remains relatively unchanged. Additionally, the above ground heights of the LiDAR returns are concentrated higher under leaf-on conditions, especially for deciduous tree types. ANOVA results indicate that the differences between both deciduous and evergreen trees under leaf-on and leaf-off conditions are significant ($p < 0.001$) for all examined metrics.

Table 1. Comparison of Weibull parameters by tree type for leaf-off and leaf-on LiDAR datasets.

	Leaf-Off		Leaf-On	
	Deciduous	Evergreen	Deciduous	Evergreen
α mean	7.00	7.59	10.02	9.44
α std	5.05	4.85	4.25	4.28
α median	6.34	7.45	9.38	9.40
β mean	10.14	9.33	12.19	8.38
β std	15.27	13.43	12.26	11.09
β median	6.90	5.89	9.28	5.69

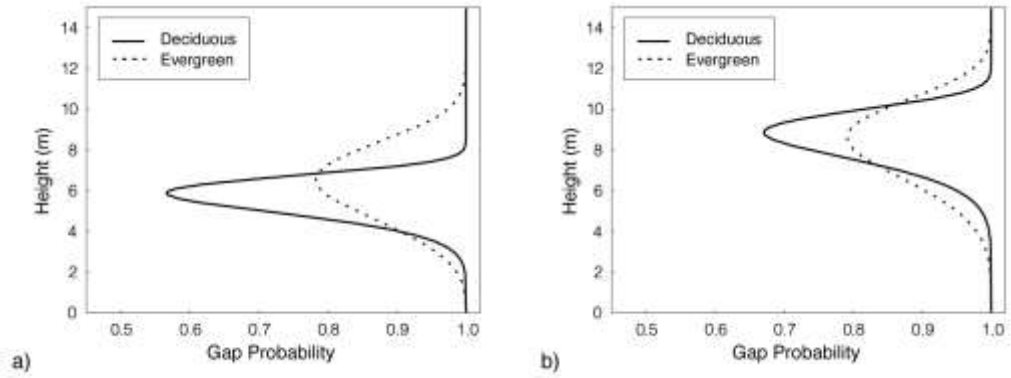


Figure 3. Vertical transmissivity profile of representative deciduous and evergreen tree types for a) leaf-off and b) leaf-off LiDAR datasets.

The modified gap probability approach was integrated with an irradiation model to assess the attenuation of radiation through the trees in the study area. Results from this procedure are shown in figure 4. Examining the source of attenuation for building walls indicates that the majority of attenuation is associated with trees on public space. A similar trend is observed for building roofs, however the split between public and private was closer across the year. Interestingly, in both cases the local attenuation from trees is generally less than 15%. The variance of the fraction of attenuation by place types is also worth examining. For building walls, variance is greatest for public space, although all places show a high standard deviation generally greater than 20%. In contrast, the variance of each place type is much reduced when looking at building roofs, generally less than 20%.

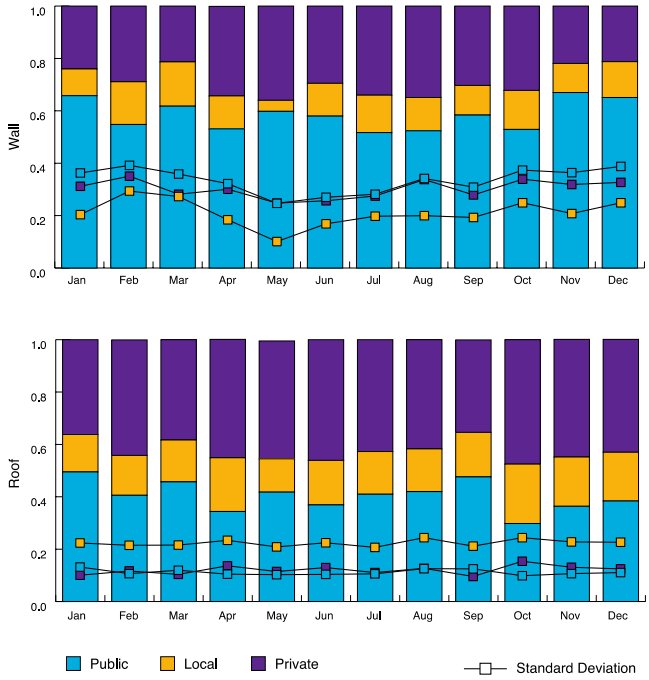


Figure 4. Radiation attenuation do to trees for building walls and roofs separated into the source space of attenuation.

4. Discussion

Results from the first part of the study, comparing basic LiDAR-based structural metrics of vegetation attenuation, suggest that statistically significant differences exist between LiDAR-derived measures for deciduous and evergreen trees and between leaf-off and leaf-on conditions. However, the dynamic nature of vegetation transmissivity, both in time and space, remains best suited to shape-based metrics for integration with common geospatial irradiance models. A technical discussion of the modified gap probability approach can be found in Tooke et al. (2012). The remainder of this section will therefore focus on the applications and implications of tree shading and irradiance modeling for energy policy.

Vegetation in the area of study included large established trees, and as a result may not be statistically representative of the entire City of Vancouver. Nonetheless, the study area does represent a common residential neighbourhood form and tree cover pattern found in many North American cities. In the introduction of this study a description was provided explaining two potentially competing building and landscape related energy policies. In the context of our study area, the first important policy message to take away from this analysis is the relatively little influence a property owner has in controlling shade/irradiance through vegetation management on their property. In the case of building walls, although there existed substantial variance in the attenuation sources from building to building, the general trend indicates that the municipal management of street trees would have a greater impact on both energy conservation opportunities and the potential generation of energy from rooftop solar technologies.

In terms of building codes and standards, the aforementioned findings demonstrate the potential inefficiency of local shading recommendations in the context of a building's broader environmental conditions. Solar access policy, on the other hand, remains important for consideration, although easements may also require acknowledgement of publically owned and operated amenities (i.e. street trees). At the same time, street trees may be providing a valuable shading service with direct implications

for building cooling requirements and the overall energy landscape of a city. In response, thinning or cutting initiatives are likely to have noticeable consequences on the energy balance of buildings.

Results of tree shade simulation studies often neglect the environmental conditions outside the immediate vicinity of a building (Pandit & Laband, 2010). As a result, tree shade policy decisions for energy management should be made aware of the generalizations that may result in ineffective outcomes. The approach presented in this paper provides a location specific methodology that can be used to better inform such policy decisions. While the analysis here focused on a treed residential neighbourhood, a similar approach is recommended for different patterns of form and tree cover to expand the discussion of policy implications. This study also helps to highlight the potential insight advanced spatial data and technology can provide policy and decision-making processes, which may be more difficult to assess using context-unspecific simulations.

5. Conclusion

In this paper we present an approach to integrating the dynamic nature of vegetation radiation transmission into current geospatial irradiance models and then apply this approach with a discussion of tree shading policy. Our results help to highlight the potentially contrasting nature of shading policies, while proposing the use of advanced spatial data, such as LiDAR, to help better inform future local energy policy design. The modified gap probability approach is well-suited for integration with existing irradiance modeling and as a result can be readily applied in a planning or policy context. In our analysis of a treed residential neighborhood, we demonstrate that the source of most radiation attenuation comes from trees located on public or private lots. The consequences of these results have important implications when informing energy policy. We suggest that solar access and building energy standards are two policies requiring consideration of the broader shading context, and can be better informed using LiDAR-based assessments of shade, such as that presented here.

References

Akbari, H., 2002. Shade trees reduce building energy use and CO2 emissions from power plants. *Environmental Pollution,* 116, 119-126.

Coops, N, Hilker, T., Wulder, M., St-Onge, B., Newnham, G., Siggins, A., Trofymow, J., 2007. Estimating canopy structure of Douglas-fir forest stands from discrete-return LiDAR. Trees-Structure and Function, 21, 295–310.

Gueymard, C., Myers, L., 2008 Solar radiation measurement: progress in radiometry for improved modeling. In: Badescu V (Ed.) *Modeling solar radiation at the earth surface*. Heidelberg: Spring-Verlag Berlin Heidelberg: 1-25.

Huang, Y., Akbari, H., Taha, H., Rosenfeld, A., 1987. The potential of vegetation in reducing summer cooling loads in residential buildings. *Journal of Applied Meteorology*, 26, 1103–1116.

Kumar, L., Skidmore, A., & Knowles, E., 1997. Modelling topographic variation in solar radiation in a GIS environment. *International Journal of Geographical Information Science*, 11, 475-497.

Laband, D., Sophocleus, J., 2009. An experimental analysis of the impact of tree shade on electricity consumption. *Arboriculture and Urban Forestry,* 35, 197–202.

Lovell, J., Jupp, D., Culvenor, D., Coops, N., 2003. Using airborne and ground-based ranging lidar

to measure canopy structure in Australian forests. *Canadian Journal of Remote Sensing*, 29, 607-622.

Ni-Meister, W., Jupp, D., Dubayah, R., 2001. Modelling lidar waveforms in heterogeneous and discrete canopies. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1943-1958.

Oke, T., 1982. The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 108, 1-24.

Pandit, R., Laband, D., 2010. Energy Saving from Tree shade. *Ecological Economics*, 69, 1324-1329.

Riaño, D., Chuvieco, E., Condés, S., González-Matesanz, J., & Ustin, S., 2004. Generation of crown bulk density for Pinus sylvestris L. from lidar. *Remote Sensing of Environment*, 92, 345-352.

Riaño, D., Meier, E., Allgöwer, B., Chuvieco, E., & Ustin, S., 2003. Modeling airborne laser scanning data for the spatial generation of critical forest parameters in fire behavior modeling. *Remote Sensing of Environment*, 86, 177–186.

Rich, P., Dubayah, R., Hetrick, W., & Saving, S., 1994. Using viewshed models to calculate intercepted solar radiation: applications in ecology. *American Society for Photogrammetry and Remote Sensing Technical Papers*, 524-529

Simpson, J., McPherson, E, 1996. Potential of tree shade for reducing residential energy use in California. *Journal of Arboriculture*, 22, 10–18.

Tooke, T., Coops, N., Christen, A., Gurtuna, O., Prevot, A., 2012. Integrated irradiance modelling in the urban environment based on remotely sensed data. *Solar Energy*, available online at: http://dx.doi.org/10.1016/j.solener.2012.06.026.

Mapping vegetation structure in a wooded savanna at Freeman Ranch, TX using airborne waveform lidar

Amy Neuenschwander
Applied Research Laboratories
University of Texas at Austin
amyn@arlut.utexas.edu

Paper Number: ### SL2012-029

Abstract

Small footprint, full-waveform lidar can provide a unique ability to characterize the landscape. The returned laser waveforms indicate specific reflectors within the footprint (vertical structure), while the shape of the return is influenced by the surface reflectance and physical topography. Recording the entire reflected waveform provides a detailed analysis of distributed targets/features along the laser line of sight. Recording of the full-waveform is especially effective in vegetative regions with respect to canopy structure characterization. The Freeman Ranch, located near San Marcos, Texas and operated by Texas State University, is a research area 4204 acres in size and is a mixture of rangeland and woodlands. The Freeman Ranch (29° 56'N, 98° W) lies within the Balcones Canyonland subregion of the Edwards Plateau and this region is undergoing successional change from grassland to Oak (*Quercus virginiana*)-Juniper (*Juniperus ashei*) dominated woodlands. The non-homogeneous landcover of Freeman Ranch, a mixture of rangeland and woodland, provides an invaluable resource for ecosystem investigations and offers an opportunity to discriminate various types of biomes and vegetative species based on their characteristic response to full-waveform lidar. Full waveform data were collected over Freeman Ranch in October 2010 and parameters such as amplitude, pulse width, integrated canopy energy, and rise time were computed for each waveform. The parameters derived from each lidar waveform were referenced to a regular grid using simple averaging for each 1m grid cell. Thus, individual data layers were developed for the amplitude of each peak component, the pulse width of each peak component, canopy energy, rise time, vertical distribution ratio, and vegetation height. These parameters were found to discriminate different vegetation morphologies within the wooded savanna that are not discernible using passive optical remote sensing data (i.e. Quickbird). Knowledge of vegetation structure such as canopy cover, vertical energy distribution, landcover type, and underlying topography improve the success of ecosystem applications and modeling.

1. Introduction

The use of remotely sensed data for landcover classification has been an active research topic since these data were made available to the science community. A limitation of classifying landcover over dense vegetation using passive multispectral or hyperspectral data is the detected reflectance is often from the upper canopy and information regarding the vertical distribution of canopy structure is not recorded. The potential of airborne lidar data for the characterization of man-made (buildings) and natural surfaces is possible due to the 3d nature of the data. In several instances the height information derived from lidar is used as ancillary data for landcover classification (Rottensteiner et al., 2004; Koetz et al., 2008). There has been considerable research on the use of large-footprint airborne waveform-sampling lidar to estimate common indices of vegetation structure including above ground biomass, basal area, stem density (Means

et al., 1999; Dubayah and Drake, 2000; Drake et al., 2002; Hyde et al., 2005; Rosette et al., 2009; van Leeuwen and Nieuwenhuis, 2010) in biomes including temperate coniferous forests in the Pacific Northwest and temperate deciduous forests in the eastern United States (Lefsky et al., 1999; Parker et al., 2001, 2004; Andersen et al., 2006). In these studies, the laser footprint size ranged from 10 – 25 m in diameter and extracted height metrics often used are RH100 (or height associated with the first reflecting surface which is inferred as top of the canopy) and HOME (or height associated with the median of the cumulative energy above the noise background) (Drake et al., 2002). The experience and techniques developed with large footprint airborne sensors (such as LVIS and SLICER) has been applied to data collected by the GLAS (Geoscience Laser Altimeter System) instrument onboard ICESat. Although the ICESat/GLAS footprint size is large in diameter, several studies indicate that heights derived from ICESat/GLAS waveforms were useful indicators of forest basal area, timber volume, and biomass (Lefsky et al., 2007; Simard et al., 2007; Rosette et al., 2008; Nelson et al., 2009; Popescu et al., 2011). In addition to height metrics, other metrics such as slope of the leading edge and total waveform extent derived from laser waveforms were found to be useful for predicting biomass (Boudreau et al, 2008; Rosette et al., 2008). Similarly, metrics such as leading edge slope, HOME, number of Gaussians, peak amplitude, and area under the 3rd Gaussian peak have been used as predictors of timber volume (Nelson et al., 2009).

Recently, small-footprint full-waveform systems are entering the commercial sector and can potentially provide a wealth of information beyond an XYZ point cloud, which is the typical commercial discrete return lidar product. For full-waveform, as the name implies, the entire temporal profile of the transmitted *and* reflected energy is recorded and stored for further processing, while the discrete return systems allow for some small number of range measurements among the reflected optical energy per laser output. Thus, full-waveform lidar offers the ability to determine additional surface information beyond what is potentially available from discrete return or photon counting systems at a high spatial resolution. Like large footprint systems, the returned laser waveforms from small-footprint laser systems indicate specific reflections within the footprint (vertical structure), while the shape of the waveform is influenced by the surface reflectance, surface roughness, and physical topography. Recording of the full-waveform is especially effective in vegetative regions with respect to canopy structure characterization. Preliminary research of waveform derived parameters has shown success in characterizing surface classes (Neuenschwander et al., 2009) as well as classifying vegetation. Knowledge of vegetative structure such as canopy cover, vertical energy distribution, and underlying topography yields improved landcover mapping and subsequently increases the success of applications ranging from ecosystem modeling to habitat suitability. The objective of this paper is to assess the potential of extracted waveform metrics beyond height metrics for describing vegetation structure and classifying the landcover within a wooded savanna. Prediction of landcover type based on structural differences in the vegetation on the landscape is a critical component for better understanding ecosystem function.

2. Method

2.1 Study area

The Freeman Ranch, located near San Marcos, Texas and operated by Texas State University, is a research area 4204 acres in size and is a mixture of rangeland and woodlands (Figure 1). The Freeman Ranch (29° 56'N, 98° W) lies within the Balcones Canyonland subregion of the Edwards Plateau and this region is undergoing successional change from grassland to Live Oak (*Quercus virginiana*)-Ashe Juniper (*Juniperus ashei*) dominated woodlands (Heilman et al., 2009). The vegetation at Freeman Ranch represents ecosystems on the Edwards Plateau, a 93,000 km^2 karst ecoregion in central Texas where woody encroachment (i.e. Ashe Juniper) has become a problem. The soil at Freeman Ranch is Comfort series, which are shallow to indurated

limestone bedrock and formed in a clayey residue. Due to the high limestone content, the soils have limited moisture storage and this area is water limited (Bendevis et al., 2010). Field data have been collected at Freeman Ranch on a quarterly basis since 2004 as part of an overall vegetation inventory effort. Within the woodlands, the Live Oak trees have canopy heights ranging from 9 – 13 m with an average diameter at breast height (dbh) of 42 cm and often have a dense understory of shrubs and Juniper beneath them. Ashe Juniper heights approach heights of 6 m in open areas and 7 – 9 m in height in dense woodlands and have an average dbh of 19 cm. Cedar Elm (*Ulmus crassifolia*) is also present in the woodland vegetation, particularly in low-lying ephemeral streams and canopy heights range from 9 – 14 m and the main stem has an average dbh of 31 cm. The heterogeneous landcover of Freeman Ranch, a mixture of rangeland and woodland, provides an invaluable resource for ecosystem investigations and offers an opportunity to discriminate various types of biomes and vegetative species based on their characteristic response to full-waveform lidar.

Figure 1. Oak/Juniper wooded savanna at Freeman Ranch, Texas

2.2 Waveform Data Analysis

Small-footprint, full-waveform lidar data were collected over Freeman Ranch in October 2010 by the National Center for Airborne Laser Mapping (NCALM) to better characterize central Texas ecosystems. NCALM's small-footprint, full-waveform lidar instrument is an Optech Gemini with an integrated waveform digitizer. The system utilizes a Nd:YAG laser operating in the near-infrared (1064 nm) with a laser pulse repetition rate of 75 kHz. The waveform digitizer samples the outgoing and return pulse at a rate of 1 ns which provides a ranging resolution of approximately 15 cm. A narrow beam divergence (0.2 mrad) was used during the acquisition and at the flight altitude the resulting footprint diameter size is roughly 15 cm. The survey was flown such that the average point density was 9 points per m^2. The raw waveforms and the aircraft attitude information were provided to the University of Texas for processing and waveform analysis. In all, 9 km^2 over Freeman Ranch were mapped during the October 2010 data acquisition.

The feature extraction methodology and classification for the full-waveform lidar data utilized in this research are designed with the desire to reduce the amount of post-processing time yet provide valuable information to characterize the canopy and canopy structure. A number of waveform features, as depicted in Figure 2, such as total integrated waveform energy and rise time of the leading edge can be extracted directly from the raw waveform. In addition to these raw waveform metrics, specific waveform features from individual scatterers or reflecting surfaces can be defined following decomposition of the lidar data into individual Gaussian components such as amplitude and pulse width. Each waveform is individually analyzed and processed such that each Gaussian peak identified within the waveform is geolocated. The standard deviation of a peak component is a potential indicator of the range distribution of scatterers, surface roughness and/or surface slope as pulse broadening occurs under these conditions.

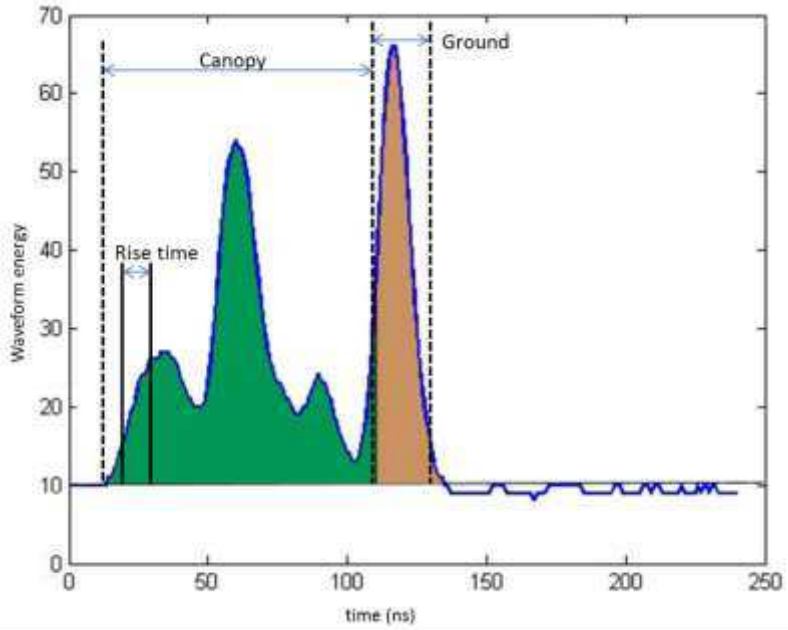


Figure 2. Illustration of small-footprint laser waveform. In this example, four distinct peaks are apparent over the background noise. The rise time is computed on the leading edge of the waveform. The area under the first three peaks represents canopy (shaded green) and the area under the last return represents the presumed ground return (shaded brown). An independent test of the last return elevation against a digital terrain model is conducted to determine whether this partitioning between canopy and ground is valid.

Other computed waveform metrics include integrated canopy energy, integrated total waveform energy, rise time to the first peak, canopy height, Height of Median Energy (HOME), and a vertical distribution ratio (VDR). Canopy, in the waveform context, refers to those parameters associated with portions of the waveform occurring before the initiation of the presumed ground reflection –in this case the last peak. However, for small-footprint systems, the assumption that the last return is the ground is not always correct. In dense canopy, the last return from small-footprint systems can often be attributed to returns from within the canopy. To ensure whether a detected last return corresponds to the ground, a separate digital terrain model with a 1 m grid posting was used in the waveform processing. If the elevation of the last return fell within half a meter of the gridded terrain elevation, that peak was assigned as a ground return. The ground return starting location is defined as the leading edge full-width-half-maximum (FWHM) position of the ground peak for a given laser shot. Canopy energy refers to the integrated waveform energy from the initial leading edge of the waveform to the beginning of the ground return. If all the detected peaks within a waveform are associated with canopy returns, then the canopy energy is the integrated energy for the entire waveform. The integrated energy parameters used for canopy and ground waveforms both provide an indication of the surface reflectivity response as they combine amplitude and pulse shape in one specific feature. The rise time is computed as the time duration between the 10 – 90% energy of the leading edge of the waveform. In the literature, the rise time is described as being related to the amount of canopy openness or variance of the upper canopy (Carabajal and Harding, 2006). For large footprint systems, vegetation height could be directly computed as the height between the last detected peak (presumed to be the ground) and first detected peak. However, again for small-footprint systems the assumption that the last return is the ground is not always true. Here, canopy height is defined as the height difference between the first return on the waveform and the elevation from the closest digital terrain surface grid cell. Similarly, Height of Median Energy (HOME) is the height associated with the median energy location along a cumulative distribution of the laser energy through the canopy above the digital terrain model.

The parameters derived from each lidar waveform were referenced to a regular grid using simple averaging for each 1m grid cell. Thus, individual data layers were developed for each waveform parameter and include the amplitude of each Gaussian peak component, the standard deviation of each Gaussian peak component, integrated canopy energy, total waveform energy, rise time, and vegetation height. Since this research is focused on using the extracted waveform parameters for classifying vegetation and canopy structure, only relevant Gaussian peaks were used to create a gridded product. When waveforms had two or more peaks, only the peaks corresponding to the canopy were considered when computing the canopy energy, pulse width, and canopy amplitude. Thus, returns from the ground did not influence the canopy level calculations. The vertical distribution ratio (VDR) is computed as

$$VDR = \frac{CH - HOME}{CH} \quad (1)$$

where CH is the canopy height. The VDR was computed at the raster level rather than at a per waveform/shot level. Ultimately, each of these direct and computed waveform features offer the capability to classify the landcover and characterize the vegetation structure based on the interaction of surface and laser energy.

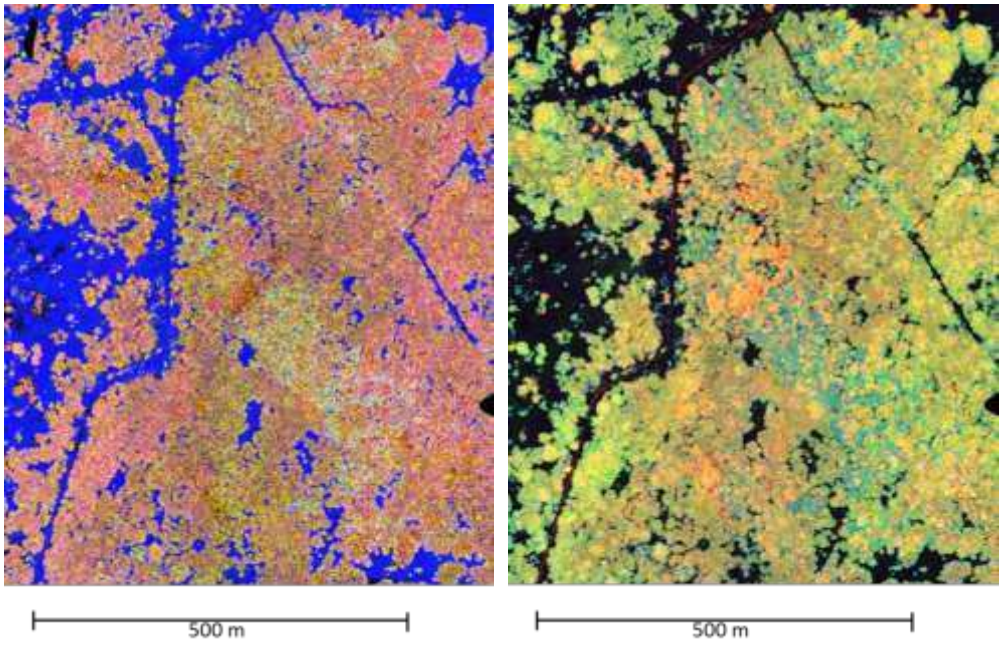


Figure 3a. False color composite of Oak/Juniper woodland savanna at Freeman Ranch using waveform parameters canopy energy, pulse width, and amplitude 3b) False color composite of Oak/Juniper woodland savanna at Freeman Ranch using waveform parameters canopy height, canopy energy, and rise time

The advantage of waveform lidar over discrete return systems is that the returned waveform can potentially detect subtleties of canopy structure that are not discernible with discrete return lidar data. Two subset images (650 x 750 m) shown in Figure 3a and 3b reveal structural/morphological differences in the Oak/Juniper woodland vegetation found at Freeman Ranch. Figure 3a is a false color composite of canopy energy, pulse width, and amplitude. The amplitude of the herbaceous layer is brighter than woody vegetation and those regions are easily recognized as blue. Although the vegetation at Freeman Ranch is generally considered an

Oak-Juniper woodland, there are patches within this mosaic that are discernible in the waveform lidar imagery that likely correspond to patches dominated by Oak, Juniper or Cedar Elm. For example, subtle differences in Figure 3a of the woody vegetation (shades of pink to green) reflect differences in pulse broadening and amplitude. Figure 3b is a composite of canopy height, canopy energy, and rise time and reveal other differences in the vegetation structure. Here, the herbaceous layer appears black due to low canopy height, low canopy energy, and low rise time. In particular, pockets of Cedar Elm woodland are visible as bright orange trees. The pixels shaded blue correspond to areas of Juniper regrowth or encroachment and these trees are short and immature.

3. Results

Five landcover classes based on species and morphological differences were identified for the study area and the classes include: herbaceous, Oak dominated woodland, Juniper dominated woodland, young juniper woodland, and Cedar Elm dominated woodland. The herbaceous layer is comprised of both native and invasive grasses ranging from 0.1 – 1 m in height. During the lidar acquisition, the herbaceous layer was approximately 50 cm in height. As discussed previously, the woody vegetation at Freeman Ranch is a mosaic of Oak/Juniper woodland with the majority of differences between each class is the dominance of one species within the mosaic and a presence of understory validated by extensive field campaigns since 2004. For example, the Oak Woodland class is typically characterized as having an Oak over-story with Juniper contributing to the understory and mid-story vegetation. The Juniper class is predominantly Juniper vegetation at mid-story and over-story levels with a clear understory. The new Juniper class consists of new Juniper vegetation that has encroached into open spaces and has canopy heights typically less than 4 m and an average dbh of 10 cm. Finally, the Cedar Elm woodland class is dominated by Cedar Elm and is typically found within the drainage channels on the property. Structurally, Oak and Cedar Elm look similar in the field and can only be correctly identified at the leaf level. The Cedar Elm class, however, can be characterized as having significantly less understory and mid-story vegetation than the Oak Woodland class. The training data for the classifier were based on polygons drawn from the waveform imagery, QuickBird imagery, and extensive field data collected quarterly since 2004. In addition to training data, additional regions of interest were digitized for independent testing of the classification results.

Table 1. Statistics (mean/st.dev) of waveform lidar metrics for five vegetation classes

	Herbaceous	Oak Woodland	Juniper Woodland	New Juniper	Cedar Elm Woodland
Canopy Height (m)	0.3 / 0.1	9.1/1.7	6.7/1	3.6/0.7	10.6/1.6
Pulse Width (ns)	1.9 / 0.2	2.5/0.3	2.5/0.2	2.9/0.4	2.5/0.3
Canopy Energy	238/77	1342/141	1050/107	1083/140	1097/135
Amplitude	223 / 24	146/27	133/22	151/19	121/34
Rise Time	3.5/.4	4.1/.4	4.1/.4	4.6/0.7	4.0/.4

The mean and standard deviation of the extracted waveform parameters were computed for each class and are listed in Table 1. The mean values for the different classes provide a sense of the separability between landcover classes based on these waveform lidar parameters. For example, the average peak amplitude for herbaceous layer is 50% higher as that for woody vegetation. The combination of the canopy energy and canopy height is perhaps the best parameters for

discriminating the woody vegetation classes based on the different morphologies for these classes. The Cedar Elm woodland class, which is characterized by less understory and midstory vegetation, has less canopy energy than the Oak Woodland class which tends to have more understory vegetation.

A support vector machine (SVM) classifier using a radial basis function as the kernel was utilized on the waveform parameters to produce a landcover classification map. The input features to the classifier included canopy height, pulse width, canopy energy, rise time, and amplitude. The thematic output from the SVM is shown in Figure 4 and an accuracy matrix is provided in Table 2. An independent test data set was used for the classification accuracy assessment. The overall classification accuracy was found to be 91.8% (kappa = 0.896) and the accuracy results for individual classes are listed in Table 2.

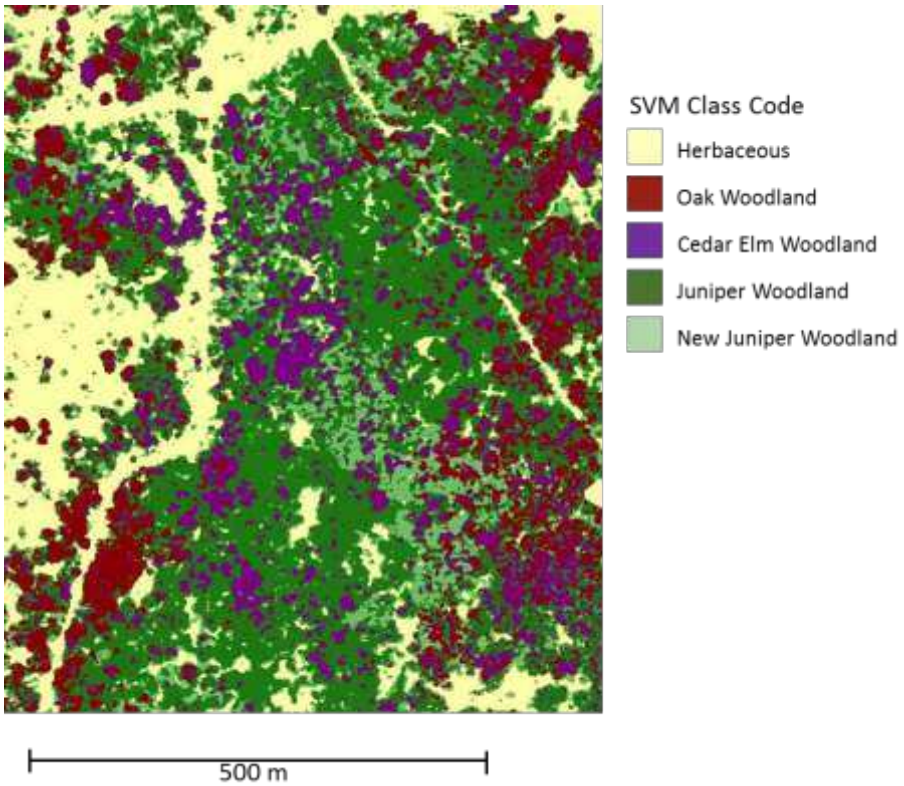


Figure 4. Support Vector Machine classification output of waveform data at Freeman Ranch

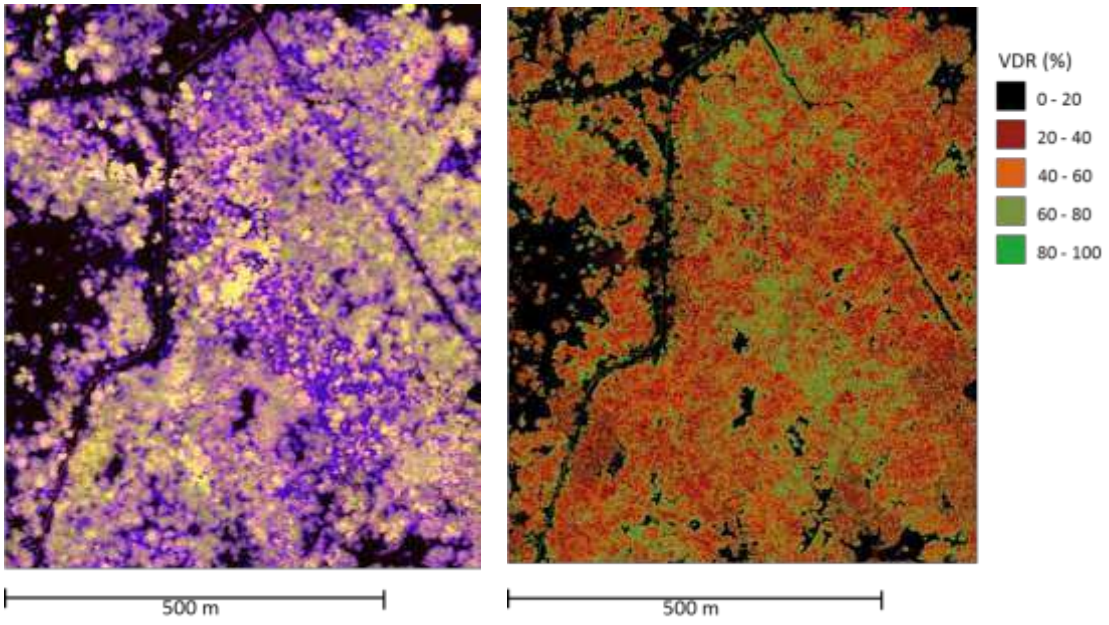
Overall, the classification accuracy using the waveform parameters is excellent at 92%. The herbaceous vegetation was easily separable from the various woodland classes as evidenced in the imagery shown in Figure 3a and 3b. The majority of confusion occurred between the Oak Woodland class and the Cedar Elm class. This confusion, however, this is not unexpected since the Cedar Elm trees look similar to the Oak: it is just lacking some of the understory vegetation found in the Oak class. Although the vegetation at Freeman Ranch is regarded as an Oak/Juniper woodland, waveform parameters reveal subtle differences in vegetation structure, i.e. morphological differences, which are invaluable for classifying the landcover at a high spatial resolution. Parameters derived from laser waveforms (i.e. amplitude, canopy energy, pulse width, etc.) reflect morphological differences between these different classes.

Table 2. Accuracy Matrix of SVM classification results, in pixels.

	Grass	Oak	Juniper	New Juniper	Cedar Elm	**USERS pixels**	**USERS %**
Grass	569	0	0	0	0	**569**	**100.0**
Oak Woodland	0	417	0	0	0	**417**	**100.0**
Juniper Woodland	0	2	409	15	5	**431**	**94.9**
Young Juniper	0	0	0	215	0	**215**	**100.0**
Cedar Elm	0	167	0	0	525	**692**	**75.9**
PRODUCERS pixels	**569**	**586**	**409**	**230**	**530**		
PRODUCERS %	**100.0**	**71.2**	**100.0**	**93.5**	**99.1**		**91.86**

Total: 91.86% Kappa coefficient: 0.896

In the literature, height metrics such as HOME and canopy height have been shown to be related to biomass and stem density via allometric regression models (Drake et al., 2002; Lefsky et al., 2002) and it is questioned whether these same parameters provide sufficient separability of the same landcover classes. For this study, VDR was computed by calculating the normalized difference between canopy height and HOME at the raster level. Figure 5a is a false color composite of these three height metrics with canopy height, HOME, and VDR. Figure 5b depicts the VDR grouped into five density classes. The pixels highlighted in red and orange have VDR values ranging from 0.2 – 0.6 indicating the majority of the reflected energy is from the upper portion of the canopy. Pixels colored shades of green have VDR values ranging from 0.6 – 1 and indicate that the energy is more evenly distributed throughout the entire height profile of the vegetation. From an ecological perspective, VDR can be used to indicate which portions of the landscape have more or less understory which could identify the potential availability of ecological niches.



(a) (b)

Figure 5. (a) False color composite of height metrics: canopy height, HOME, and VDR. (4b) Color coded map of vertical distribution ratio (VDR).

A SVM classifier using the same training and testing data as previously described was run on height metric inputs (i.e. canopy height, HOME, and VDR) to produce a landcover classification map. The thematic output from the SVM is shown in Figure 6 and an accuracy matrix is provided in Table 3. The thematic output by including only these three input features resulted in misclassification and confusion between the Oak woodland and Cedar Elm woodland classes. Specifically, only one pixel of the Oak Woodland class was correctly identified. These results demonstrate the importance of additional waveform parameters such as integrated canopy energy, amplitude, and pulse width to characterize and classify the vegetation.

Table 3. Accuracy Matrix of SVM classification results using height metrics only, in pixels.

	Grass	Oak	Juniper	New Juniper	Cedar Elm	**USERS pixels**	**USERS %**
Grass	569	0	0	1	0	**570**	**99.8**
Oak Woodland	0	1	4	0	0	**5**	**20.0**
Juniper Woodland	0	9	399	32	5	**440**	**90.7**
Young Juniper	0	0	6	197	0	**203**	**97.0**
Cedar Elm	0	576	0	0	530	**1106**	**47.9**
PRODUCERS pixels	**569**	**586**	**409**	**230**	**530**		
PRODUCERS %	**100.0**	**0.2**	**97.5**	**85.6**	**100.0**		**72.9**

Total: 72.97% Kappa coefficient: 0.657

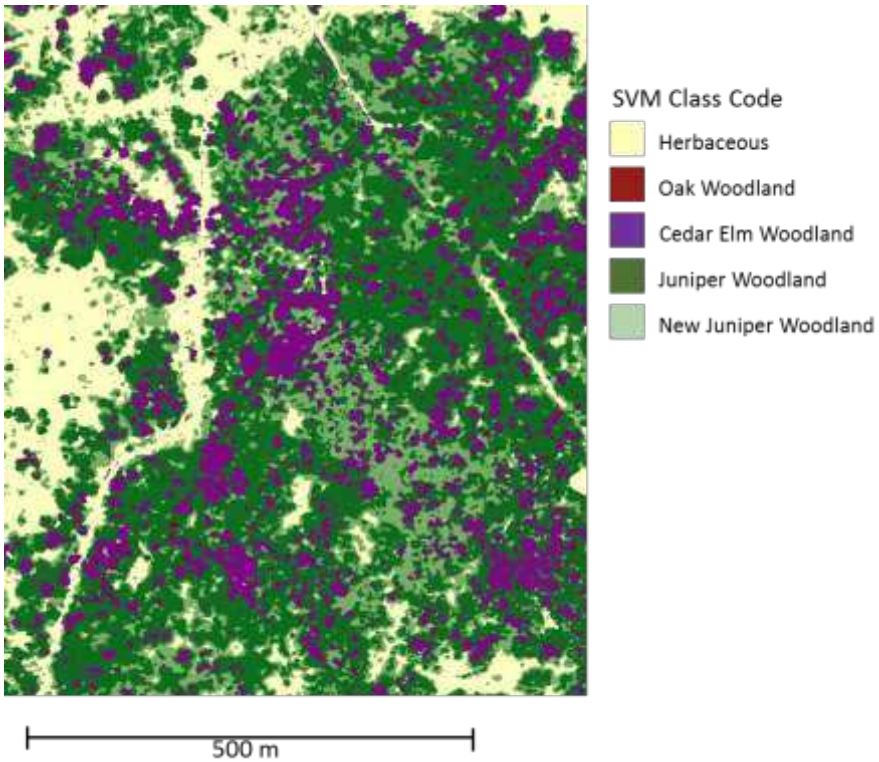


Figure 6. Thematic output from SVM classification using on lidar height metrics as inputs.

4. Discussion

The results when using waveform lidar parameters to classify landcover based on structural or morphological differences within the canopy are encouraging; however, the scope of this study was limited to a wooded savanna ecosystem in central Texas. Traditionally, waveform lidar parameters such as canopy height, slope of the leading edge, and HOME are used as variables to predict biomass or other canopy structure variables (Drake et al., 2002; Nelson et al, 2009). These studies often employ different allometric models based on landcover typically derived from other sensors (e.g. MODIS or Landsat) or stratify the output based on landcover type. The research presented here is a novel application of waveform lidar data to improve characterization of terrestrial ecosystems. The ability to classify landcover directly from the waveform lidar data based on morphological differences in the vegetation structure shows potential as an effective utilization of this data. This ability is important for areas that are spatially heterogeneous and therefore too spectrally mixed in optical imagery. This technique could be applied and tested for other similar ecosystems with homogeneous pockets of vegetation.

In this study, lidar waveform parameters such as amplitude, pulse width, rise time, canopy height, and integrated canopy energy resulted in a 92% landcover classification accuracy. The same methodology, however, could be applied to other ecosystem types to ascertain as to whether the same waveform parameters are useful. Since height and other height metrics are relatively easy to derive from lidar data, a similar landcover classification was implemented using only height metrics (canopy height, HOME, and VDR) and the results were less successful. In particular, the Oak woodland was completely misclassified as Cedar Elm woodland. This misclassification indicates that additional waveform lidar parameters such as canopy energy, amplitude, and pulse width are important to discriminate these morphological classes. An effective utilization of lidar waveforms for landcover classification relies upon different morphologies of the vegetation as this information is inherently what is recorded in the waveform data.

Analysis of large-footprint, airborne and space-based laser waveforms has been an active research topic for the past decade and most recently, small-footprint full-waveform laser mapping systems are entering the commercial market which can provide additional detail beyond discrete return systems. Commercial vendors such as Optech and Reigl offer full-waveform digitization capabilities with the new laser systems for sale. Over vegetation, full-waveform lidar data provide a detailed description of canopy structure along the laser line-of-sight since the returned laser waveform consists of the reflected laser energy through the canopy. Large-footprint waveforms are not typically utilized for classifying landcover, however, this research has shown that waveform lidar metrics from small-footprint systems are useful for classifying landcover and characterizing the terrain. This capability to map landcover based on different vegetation morphologies is particularly relevant for ecologists or natural resource managers who may originally be required to collect two lidar surveys: one at leaf-off to discriminate evergreen from deciduous vegetation using discrete return lidar and again at leaf-on to map the crown area and canopy height for all vegetation types. Using small-footprint waveform lidar, the potential to map canopy and landcover based on structure or vertical configuration from only one lidar survey opens new possibilities for ecological mapping applications and saves money.

A research area that should be addressed in future work is to determine if there are other waveform metrics that can be derived from small-footprint waveform lidar that can improve estimates of vegetation structure. One challenge with using waveform derived parameters is

their direct interpretability with vegetation that can be measured and validated in the field at a fine spatial scale. For example, is the rise time of the leading edge correlated with crown closure or penetration into the canopy? In previous research, the energy penetration index (EPI) which is a ratio of ground energy to total waveform energy was shown to be related to canopy cover for large footprint systems (i.e. ICESat/GLAS) and is a good indicator for canopy characterization (Neuenschwander et al., 2008; Popescu et al, 2011). However, at a smaller footprint size the physical interpretation of the waveform data are different than with laser waveforms collected from large footprint systems. For example, a small-footprint waveform with a footprint diameter of 20 cm will only describe a portion of the tree structure. It is suggested that fine-scale waveform data be aggregated to the tree-level to compute tree-level statistics, which can then be validated with field derived measurements, while the within-tree variability can still provide subtle discrimination capabilities for different vegetation types.

Acknowledgements

This research was funded under the NASA Carbon Cycle Science Program NNX11AG91G, PI Marcy Litvak at the University of New Mexico. Many thanks to National Center for Airborne Laser Mapping (NCALM) for providing the waveform lidar data for this research. Thanks also go to Dr. Joel P. Bach at Texas State University for access to Freeman Ranch and support of this research. Thanks also to the reviewers of this paper for their feedback and comments.

References

Anderson, J., Martin, M.E., Smith, M-L., Dubayah, R.O., Hofton, M.A., Hyde, P., Peterson, B.E., Blair, J.B., and Knox, R.G., 2006. The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sensing of Environment*, 105: 248-261.

Bendevis, M.A., Owens, K., Heilman, J.L., and McInnes, K.J. 2010. Carbon exchange and water loss from two evergreen trees in a semiarid woodland. *Ecohydrology*, 3(1):107-115.

Boudreau, J., Nelson, R., Margolis, H., Beaudoin, A., Guindon, L., Kimes, D., 2008. Regional aboveground forest biomass using airborne and spaceborne lidar in Quebec. *Remote Sensing of Environment*, 112: 3876-3890.

Carabajal, C.C. and Harding, D.J. 2006. SRTM C-band and ICESat Laser Altimetry Elevation Comparisons as a Function of Tree Cover and Relief. *Photogrammetric Engineering and Remote Sensing* 72(3), 287-298.

Dubayah, R. and Drake, J., 2000. Lidar remote sensing for forestry. *J. Forestry*, 98(6):44-46.

Drake, J.B., Dubayah, R.O., Clark, D.B., Knox, R.G., Blair, J.B., Hofton, M.A., Chazdon, R.L., Weishampel, J.F., and Prince, S.D., 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, 79: 305-319.

Heilman, J.L., McInnes, K.J., Kjelgaard, J.F., Owens, M.K., and Schwinning, S., 2009. Energy balance and water use in a subtropical karst woodland on the Edwards Plateau, Texas. *J. of Hydrology*, 373:426-435.

Hyde, P., Dubayah, R., Peterson, B., Blair, J.B., Hofton, M., Hunsaker, C., Knox, R., and Walker, W., 2005. Mapping forest structure for wildlife habitat analysis using waveform lidar: validation of montane ecosystems. *Remote Sensing of Environment*, 96:427-437.

Koetz, B., Morsdorf, F., Van der Linden, S., Curt, T., and Allgower, B., 2008. Multi-source land cover classification of forest fire management based on imaging spectrometry and lidar data. Forest Ecology and Management, 256(3):263-271.

Lee, S., Ni-Meister, W., Yang, W., and Chen, Q., 2011. Physically based vertical vegetation structure retrieval from ICESat data: Validation using LVIS in White Mountain National

Froest, New Hampshire, USA. *Remote Sensing of Environment*, 115: 2776-2785.

van Leeuwen, M., and Nieuwenhuis, M., 2010. Retrieval of forest structural parameters using lidar remote sensing. *European J. of Forest Research*, 129:749-770.

Lefsky, M.A., Harding, D.A., Cohen, W.B., Parker, G., and Shugart, H.H. 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of Eastern Maryland, USA. *Remote Sensing of Environment* 67, 83-98.

Lefsky, M., Keller, J., Pang, Y., deCamargo, P., and Hunter, M., 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter system waveforms. *J. Applied Remote Sensing*, 1, 013537.

Means, J., Acker, S., Harding, D., Blair, J.B., Lefsky, M.A., Cohen, W.B., Harmon, M, and McKee, W.A. 1999. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western cascades of Oregon. *Remote Sensing of Environment* 67, 298-308.

Nelson, R., Ranson, K.J., Sun, G., Kimes, D.S., Kharuk, V., and Montesano, P., 2009. Estimating siberian timber volume using MODIS and ICESat/GLAS. *Remote Sensing of Environment*, 113:691-701. [doi:10.1016/j.rse.2008.11.010]

Nelson, R., Boudreau, J., Gregoire, T., Margolis, H., Naesset, E., Gobakken, T., and Stahl, G., 2009. Estimating Quebec provincial forest resources using ICESat/GLAS. *Canadian Journal of Forest Research*, 39(4): 862-881. [doi:10.1139/x09-002]

Neuenschwander, A.L., Urban, T.J., Gutierrez, R., and Schutz, B.E., 2008. Characterization of ICESat/GLAS waveforms over terrestrial ecosystems; Implications for vegetation mapping. *J. Geophysical Research*, 113, G02S03.

Neuenschwander, A., Magruder, L., and Tyler, M., 2009. Landcover classification of small-footprint, full-waveform lidar data. *J. of Applied Remote Sensing*, 3.033644.

Parker, G.G., Lefsky, M., and Harding, D., 2001. Light transmittance in forest canopies determine using airborne laser altimetry and in-canopy quantum measurements. *Remote Sensing of Environment*: 76:298-309.

Parker, G.G., Harmon, M., Lefsky, M., Chen, J., Van Pelt, R., Weis, S., Thomas, S., Winner, W., Shaw, D., and Frankling, J. 2004. Three-dimensional structure of an old-growth Pseudotsuga-Tsuga Canopy: Its implications for radiation balance, microclimate, and gas exchange. *Ecosystems*, 7(5):440-453.

Popescu, S., Zhao, K., Neuenschwander, A., and Lin, C. 2011. Satellite lidar vs. small-footprint airborne lidar: comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. *Remote Sensing of Environment*, 115:2786-2797.

Rosette, J., North, P., and Suarez, J., 2008. Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. *Int. Journal of Remote Sensing*. 29(5):1475-1493. doi[10.1080/01431160701736380]

Rosette, J., North, P., and Suarez, J., 2009. A comparison of biophysical parameter retrieval for forestry using airborne and satellite lidar. *International Journal of Remote Sensing*. 30(19), 5229-5237.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., and Lovell, B., 2004. Building detection by Dempster-Shafer fusion of lidar data and multispectral aerial imagery. *Proceedings to International conference on pattern recognition*, 2:339-342.

Simard, M., Rivera-Monroy, V, Mancera-Pineda, J.E., Castaneda-Moya, E., and Twilley, R.R., 2007. A systematic method for 3D mapping of mangrove forests based on Shuttle Radar Topography Mission elevation data, ICESat/GLAS waveforms and field data: Application to Cienaga Grande de Santa Marta, Colombia. *Remote Sensing of Environment*, 112(5): 2131-2144 [doi:10.1016/j.rse.2007.10.012]

Automatic extraction of tree stem models from single terrestrial lidar scans in structurally heterogeneous forest environments

David Kelbe[1], Paul Romanczyk[1], Jan van Aardt[1], Kerry Cawse-Nicholson[1] & Keith Krause[2]

[1]Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY
djk2312@cis.rit.edu
[2]National Ecological Observatory Network

Keywords: terrestrial laser scanning, single scan, tree stem models, DBH

Paper ID: SL2012-031

Abstract

An important application of terrestrial laser scanning is the extraction of tree stem models for diameter at breast height (DBH) assessment and forest inventory. Much work has been done to automate this process using adjacent co-registered lidar scans. Existing studies, however, have focused on pre-registered point clouds obtained from commercial lidar systems. We envision an affordable and efficient portable lidar hardware-software alternative for forest structural assessment. For such systems, point cloud registration software may not be available. Therefore we have developed automatic tree stem and DBH modelling approaches for single lidar scans collected from an off-the-shelf SICK LMS-151 lidar system. We assessed visual and quantitative accuracy for automatic tree extraction in a dense, heterogeneous, 20m x 20m forest site. We found that modelling tapered cylinders to the lidar point distribution of tree stems presents a viable algorithm for our single-scan lidar system, even in cluttered heterogeneous forest environments. This work demonstrates the potential for forest inventory using affordable, off-the-shelf lidar systems, and offers suggestions for future research.

1. Introduction

Terrestrial laser scanning (light detection and ranging; lidar) has become an increasingly important tool for forest structure assessment in recent years. Laser scanning technology records the return trip time, from laser emission to target backscatter to final detection, of a backscattered laser pulse for a series of scan mirror positions. Range distance to the object is computed based on the speed of light, and as a result, a dense 3D point cloud is generated. This rich data source holds the possibility of extracting many forest structure parameters, including those that would be unfeasible to collect using traditional field methods (e.g., stem volume, sweep, flexuosity, and branch structure) (Gorte and Pfeifer, 2004; Othmani *et al*., 2011). Yet, due to the complexity of segmentation in forested environments, most studies so far have focused on the extraction of individual tree diameter-at-breast-height (DBH) and corresponding stem densities (Simonse *et al*., 2003; Aschoff *et al*., 2004; Maas *et al*., 2008; Othmani *et al*., 2011).

Significant advances have been made in recent years to automatically measure and extract tree DBH on an operational scale. However, most algorithms rely on expensive commercial lidar systems and associated proprietary software. Furthermore, the data acquisition time of most commercial systems prohibits large-scale inventory. We envision an affordable and efficient terrestrial lidar hardware and software solution that can be utilized for forest inventory and assessment outside the domain of research and academia. Such a product-algorithm suite could

serve as an alternative to traditional forest inventory techniques.

As such we have integrated a SICK LMS-151 laser scanning system with a rotation stage and portable computer, at a cost of under US$ 10,000. Our system allows rapid and economical forest structure data acquisition. It is our goal to draw upon the tools developed for tree structure extraction, specifically DBH, to expand our system functionality towards automatically extracting tree stem models. Specifically, we hope to investigate the extension of DBH extraction techniques to single lidar scans.

The majority of previous studies focus on modelling tree stems from multiple, pre-registered scans of a central location (Simonse *et al*., 2003; Aschoff *et al*., 2004; Maas *et al*., 2008; Othmani *et al*., 2011). Emphasis on multiple-scan algorithms has so far been motivated by the decreased measurement error due to fewer occlusions of the laser beam by understory (Watt and Donoghue, 2005). In other words, adjacent scans may provide a view of the target that was otherwise obscured in the single scan. Traditional terrestrial lidar-based DBH-approximation methods consider points from a 5cm thick horizontal z-slice centred at 1.3m above ground. A single circle is fitted to the 2D-projection of these points. Operationally, this procedure may be applied at varying heights above ground (Othmani *et al*., 2011; van Leeuwen *et al*., 2011). In contrast, relatively few studies have considered tree stem characterization from a single scan. The challenges introduced by single scans are several: First, a singular scan position causes obscuration radially outward from each intercepted object. Second, point density and resolution decrease radially due to the laser beam divergence. In one pilot study, the rate of tree detection from single scans was less than half that for multiple scans (Thies and Spiecker, 2004).

However, despite the challenges introduced from single scan algorithms, there are advantages: It is clear that a single panoramic scan is preferable economically (Thies and Spiecker, 2004). A robust single-scan algorithm could vastly increase operational efficiency for field crews. We suggest that a balance can be achieved between efficiency and accuracy, by taking advantage of the full structural information available from a single laser scan: Instead of fitting points to a circle at 1.3m, we will utilize the full set of points between 0.3m and 1.3m above ground and fit tapered cylinders. Here we present preliminary results of algorithm development for single-scan tree stem modelling by modelling tapered cylinders to the hemi-cylindrical point distribution on individual tree boles.

2. Method

2.1 Study area

Our study area is a 10ha section of the Hemlock-Canadice State Forest, located at 42.7377° N, 77.6061° W, in upstate New York (Figure 1). The Hemlock-Canadice State Forest lies adjacent to Hemlock Lake, the water source for the city of Rochester. The forest's history traces back to the early 1900's when tracts of land in the watershed of Hemlock Lake were acquired by the city to protect the public water supply. Most of the land was agricultural at the time. In order to protect the quality of the water, aggressive tree planting began shortly after. Planted species included Scots pine (*Pinus sylvestris*), white pine (*Pinus strobus*), red pine (*Pinus resinosa*), and Norway spruce (*Picea abies*), while pockets are naturally forested. In 1929, spread of a fungal disease prompted the underplanting of stressed conifers with hardwood seedlings, in order to retain forest cover should the conifers die.

2.1.1 Site establishment

We established 20 sites on the west-facing shore of Hemlock Lake, in the Hemlock-Canadice

State Forest. Site centres are separated by approximately 70-90m, depending on restrictive characteristics, such as extreme slopes at site centres. Each site is 20m x 20m in dimension. For this work a single site was selected using a pseudo-random number generator to investigate the challenges of automatically extracting tree stem models from a single lidar scan. Our site (Site #6) is a natural mixture of young deciduous and coniferous trees on variable terrain (Figure 2). Many conifer branches are present under 2m, resulting in significant visual occlusion and clutter.

Figure 1: Our site was randomly selected from 20 marked sites on the west-facing slopes of Hemlock-Canadice State Forest, located in upstate New York (a). The site is a natural mixture of deciduous and coniferous trees (b). (map: Google Maps)

2.1.2 Data

Lidar data were collected in April 2012 during the leaf-off season using the SICK LMS-151 lidar system. This system has a maximum vertical field of view of 270º, an angular step width of 0.25º and a scanning frequency of 50Hz. The instrument is mounted on a rotation stage to provide 360º field of view in azimuth at a 0.25º step width. The lidar beam diverges to give an effective beam size of 15.8cm at 10m. The system is then mounted on a tripod and tethered to a computer and power supply, which can be carried by the operator in an external frame backpack (Figure 3). This allows rapid data acquisition at the rate of 25-scans/half-hour. In addition, we measured the DBH and stem location, using a compass and tape measure from plot centre, of each tree over 5cm DBH within the site.

Figure 2: Our mobile lidar system allows rapid forest structure inventory at approximately 25 scans/half-hour.

3. Algorithm

Tree stem modelling was performed by iteratively querying a subset of points within the region between 0.4 and 1.3m above ground. Each subset of points represents a candidate stem for cylinder model fitting. A 1.3m height threshold was chosen to exclude the majority of branch and lower canopy returns. Points within the subset were then fitted to models of either lines or tapered cylinders, based on the number of points within the subset. The DBH was computed as a function of modelled parameters; the algorithm is discussed in the following sections.

3.1 DTM extraction

The digital terrain model (DTM) describes the height of the ground with respect to the lidar coordinate system origin. The DTM gives a reference from which DBH can be measured. To derive the DTM, a search window is applied to find local minima points. The search window size is radially dependent to account for decreasing point density with range. Outlier points are detected and eliminated based on a slope threshold, pre-determined from knowledge of scene terrain. The remaining points are interpolated using Delaunay triangulation. The final grid product is inpainted (Matlab; Bertalmio *et al*., 2000) to ensure DTM coverage over the entire sample area.

3.2 Line fitting

Lidar points are fit to a line if fewer than 150 points are located within the subset. In order to estimate the principal direction of the line, e.g., the tree stem growth angle, we find the direction of maximum variance using principal component analysis (PCA). The position of the stem tree centre is computed from the intersection of this vector with the DTM. The stem diameter is estimated from the standard deviation of points in the direction orthogonal to both the principal direction of growth and the line of sight vector to the lidar scanner.

3.3 Cylinder fitting

Lidar points are fitted to a tapered cylinder if more than 150 points are contained in the subset. A two-stage approach is used to a) derive an initial estimate and then b) iteratively adjust that estimate to optimize model parameters.

3.3.1 Initial Estimate

The success of optimization routines is improved when an adequate initial estimate is supplied. We derive an initial estimate for the cylinder's principal direction by calculating the direction of maximum variance using eigen analysis. The stem location is again computed as the intersection of the shifted eigenvector with the DTM. The radius is estimated from the distribution of points projected onto the vector orthogonal to both the first eigenvector and the vector to the origin. Their density distribution is fitted to a SuperGaussian with free parameters n and σ. Fitting to a SuperGaussian allows consistent stem diameter estimation for tree stems with and without branch returns. An initial estimate of the taper is fixed according to the Ormerod taper equation (Ormerod, 1973).

3.3.2 Optimization

Modelled tree stem parameters are iteratively adjusted using Trust-Region-Reflective Optimization (Coleman and Li, 1996). We establish our model as follows: The task is to minimize the distance of points from the surface of a tapered cylinder A (Figure 5), with arbitrary location (x_0, y_0, z_0), direction of growth (α,β,γ), base radius (r_0), and taper angle (θ). In order to derive a functional form for the error, we transform the points to a vertical stem centred at the origin (Figure 5, cylinder B). The transformation matrix is derived from the current estimate of (α,β,γ) and (x_0, y_0, z_0).

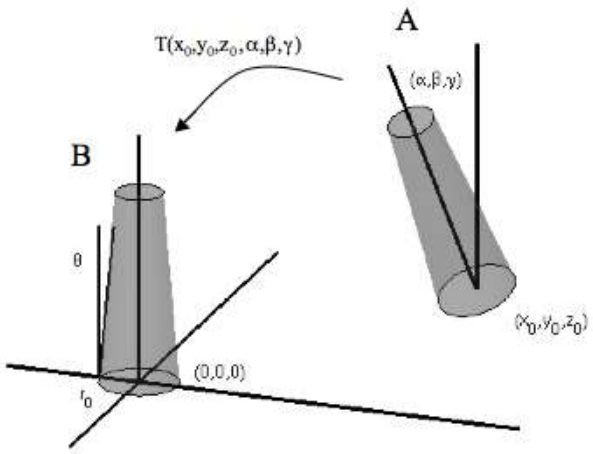


Figure 3: A schematic that shows the cylinder fitting process

The problem is then recast as minimizing

$$sum\{(x)^2 + (y)^2 - (r_0 + z * tan(\theta))^2\} \qquad (1)$$

where (x,y,z) are the observed lidar points. This process is iterated for all candidate tree stems within the lidar point cloud. The result is a list of parameterized tree stem models. Each model is defined by an (x,y,z) coordinate denoting the intersection of the tree bole with the DEM, the angular orientation or direction of growth, (α,β,γ), the taper angle (θ), and the diameter at breast height (computed). Additionally, the indices of the inlier points to the model are stored for point classification (e.g., ground, bole, etc.).

4. Results and Discussion

4.1 Visual Analysis

To visualize the results of automatic tree stem extraction, we generated a 2D panoramic image from the lidar point cloud (Figure 4). Columns correspond to sampled steps in azimuth angle, and rows correspond to sampled steps in elevation angle. The intensity of each pixel was derived from the backscattered 905nm intensity for the appropriate intercepted object.

Point classification was performed to provide a visual analysis of stem detection. Unclassified points are grey. Ground points, derived from the DTM, are blue. Stem points, derived from cylinder fitting, are orange. Note that only points used in the cylinder-fitting model (0.4m to 1.3m above ground) are classified. Good visual classification is achieved even under partial occlusion, varying terrain, branch structure, and object distance.

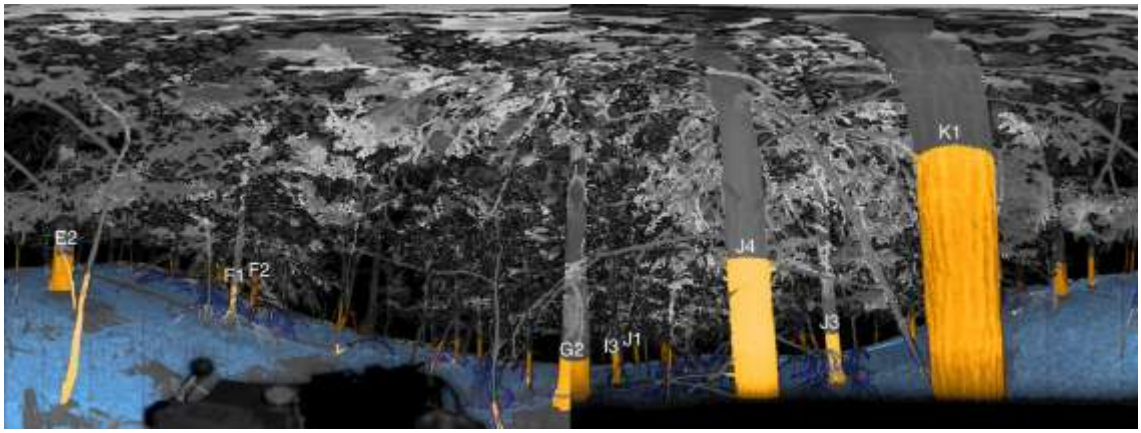


Figure 4: Point classification showing ground points (blue), stem points (orange), and unclassified points (grey), as projected onto a 2D panoramic image. Trees with >4cm DBH within a 10m x 10m site subset are labeled according to inventory number.

4.1 Quantitative Analysis

Quantitative model accuracy was evaluated by (i) direct comparison to mapped and measured tree boles within a 10m x 10m site subset, and (ii) stem diameter distribution comparison within the full 20m x 20m site. For direct comparison between boles, the DBH of each tree greater than 4cm DBH was measured and mapped using a compass and tape measure. Measured DBH values were compared to the estimated DBH values for all trees within the 10m x 10m subset. Measurement error was calculated based on the lidar beam divergence. Beam divergence introduces measurement error due to the possibility of partially intercepted objects. At a distance x from the aperture, the lidar beam diverges to a width d according to Equation 2.

$$d\,[m] = 008\,[m] + 0.015 * x\,[m] \qquad (2)$$

In addition, a simple linear regression was calculated (Equation 3) relating measured and estimated DBH values (Figure 5).

$$Estimated\ DBH = 0.071 + 0.805 * Measured\ DBH \qquad (3)$$

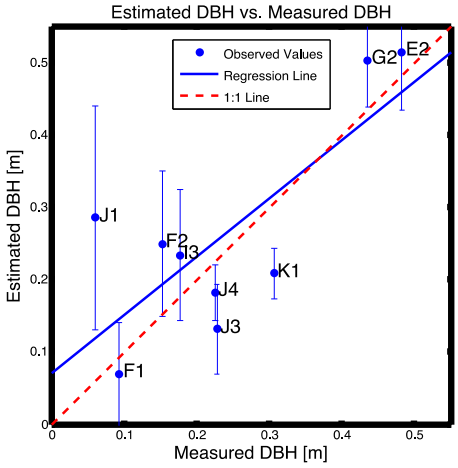


Figure 5: Direct comparison between estimated and measured DBH.

The second method for model evaluation compared stem diameter distributions across the entire 20m x 20m site (Figure 6). In other words, the distribution of inventoried tree diameters was compared to that of the estimated tree diameters, without regard for location. This permitted model evaluation across the entire site when individual tree comparison was infeasible. It is expected that the lidar-derived stem counts should be lower than the actual stem counts, due to the occlusion of objects from a single lidar scan. This phenomenon should manifest itself in an estimated diameter distribution shifted below the actual diameter distribution (i.e., fewer trees detected at each DBH bin). However, Figure 6 shows a slight overestimation of stem counts. This reveals a high false alarm rate in tree detection that was difficult to visualize in the point classification image (Figure 4). False alarms may be difficult to visualize in the point classification image if only few pixels contribute to a computed tree stem model. Therefore, future work should additionally visualize tree stem models as objects in 3D space.

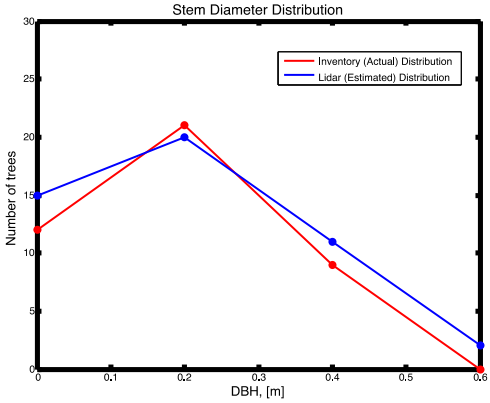


Figure 6: Stem DBH distribution comparison between estimated and measured values.

These analyses were derived from a single lidar scan collected from an affordable, off-the-shelf instrument. No proprietary pre-processing or registration software was needed, and scanning time was less than one minute. This work suggests the potential for efficient, large-scale forest inventory in the future, using economical lidar systems such as the SICK LMS-151.

5. Conclusion

An important application of terrestrial laser scanning is the extraction of tree stem models for

DBH assessment and forest inventory. We developed a method for single-scan stem detection based on existing methods for multiple-scan systems. We performed visual and quantitative assessment of automatic tree stem extraction. We found that modelling tapered cylinders to the lidar point distribution of tree stems presents a viable algorithm for our single-scan lidar system, even in cluttered heterogeneous forest environments. This has significant impact for future research, suggesting that off-the-shelf, single-scan lidar systems could extend the accessibility of forest structure assessment to cost-limited applications.

Acknowledgements

This material is based upon work supported by the National Science Foundation Accelerating Innovation Research (AIR) program and National Science Foundation Graduate Research Fellowship under Grant Nos. 1127728 and DGE-1102937, respectively. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The authors would also like to thank the NYS Department of Environmental Conservation, Region 8 for access to Hemlock-Canadice State Forest.

References

Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C., 2000. Image inpainting. In: *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (SIGGRAPH '00). ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 417-424. Freiburg, University of Friburg: 149-152.

Watt, P. J. and Donoghue, D. N. M., “Measuring forest structure with terrestrial laser scanning,” International Journal of Remote Sensing 26(7), pp. 1437–1446, 2005.

Coleman, T.F. and Y. Li, "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds," *SIAM Journal on Optimization*, vol. 6, pp. 418-445, 1996.

Othmani, A., Piboule, A., Krebs, M., Stolz, C., and Lew Yan Voon, L. (2011) Towards automated and operational forest inventories with t-lidar. Silvilaser 2011 Conference Proceedings.

Simonse, M., Aschoff, T., Spiecker, H., and Thies, M. (2003) Automatic determination of forest inventory parameters using terrestrial laserscanning. In Proceedings of the ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests, Umea, Sweden, pp. 251–257.

Maas, H., Bienert, A., Scheller, S., and Keane, E. (2008) Automatic for- est inventory parameter determination from terrestrial laser scanner data. International Journal of Remote Sensing, 29, 1579–1593.

Aschoff, T., Thies, M., and Spiecker, H. (2004) Describing forest stands using terrestrial laser-scanning. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXV of Comm. 5, pp. 237–241.

van Leeuwen, M., Coops, N. C., Newnham, G. J., Hilker, T., Culvenor, D. S., and Wulder, M. A. (2011) In: Silvilaser 2011 Conference Proceedings. Stem detection and measuring dbh using terrestrial laser scanning.

Theis, M. and Spiecker, H. (2004) Evaluation and future prospects of terrestrial laser-scanning for standardized forest inventories. In Proceedings of the ISPRS working group VIII/2 Laser-Scanners for Forest and Landscape Assessment.

Bertalmio, M., Sapiro, G., Caselles, V., and Ballester, C. (2000) Image in- painting. Proceedings of the 27th annual conference on Computer graphics and interactive techniques, New York, NY, USA, pp. 417–424, SIGGRAPH ’00, ACM Press/Addison-Wesley Publishing Co.

Ormerod, D. (1973) A simple bole model. Forestry Chronicle, 49, 136–138.

Comparing performances of ALS and Landsat 7 ETM+ satellite optical data in stratification-based sampling method for large-area forest inventory

Svetlana Saarela[1], Annika Kangas[1] Sakari Tuominen[2], Markus Holopainen[1], Juha Hyyppä[3], Mikko Vastaranta[1] & Ville Kankare[1]

[1]Department of Forest Sciences, University of Helsinki, {first.lastname}@helsinki.fi
[2]Finnish Forest Research Institute (METLA), sakari.tuominen@metla.fi
[3]Finnish Geodetic Institute, juha.hyyppa@fgi.fi

Paper Number: SL2012-032

Abstract

ETM+ satellite images have traditionally been used to provide auxiliary data for large-area forest inventories. One of their advantages is their large geographic coverage. Their disadvantages include the medium resolution and lack of 3D information. These disadvantages can be compensated through the Airborne Laser Scanning (ALS) data in combination with aerial images. Finnish Forestry Centres use Airborne Laser Scanning (ALS) data in combination with aerial images for small-scale forest inventories in the management planning of privately-owned forests. For different forest management tasks in different land-class areas, different sets of ALS features are used. The goal of this study was to investigate the performance of ALS data in combination with satellite optical data for large-area forest stratification-based inventories in order to increase the estimation accuracy. Ellipsoid stratification was used. It was found that with an increasing number of strata for stratification, the standard error for the total stem volume was found to decrease and stabilize.

1. Introduction

Forest resource information is gathered for large-scale strategic planning, operative forest management and pre-harvest planning. Examples of inventories undertaken for large-scale strategic planning are national forest inventories (NFIs) which gather information about nationwide forest resources, such as growing stock volume, forest cover, growth and yield, biomass, carbon balance and large-scale wood procurement potential. It is important that the results for the NFI provide accurate and unbiased estimates of the resource.

Laser scanning data is a far more expensive auxiliary data than, e.g., satellite images. Thus, strategic large-area forest inventories are still based either solely on field measurements (national level) or a fusion of field data and satellite images. At the county level, ALS inventory has been studied. Næsset (2004) tested area-based approach (ABA) in a 65 km^2 area in Norway. The accuracy of the predicted plot level volume was 17.5%-22.5% and the respective stand-level accuracy was 9.3%-12.2%. Holmgren and Jonson (2004) conducted a similar study in a 50 km^2 area in Sweden with a stand-level volume RMSE of 14.1%. In the aforementioned studies, ALS data covered the entire study area. In recent years, even larger areas have been inventoried operationally using ABA. At the national inventory level, it is not feasible to acquire wall-to-wall ALS data for forest inventory purposes.

Holopainen and Hyyppä (2003), and Næsset et al. (2006) suggested the use of ALS data in strip-based sampling. There have been many studies in which profiling LiDAR has been used to acquire sampled forest inventory data (e.g. Nelson et al. 2003a, 2003b, 2004). Nelson et al. (2004) inventoried forest resources in the state of Delaware in the U.S. using 14 flight lines with

a 4 km sampling distance. Their timber volume estimate at the county and state level differed from the U.S. Forest Service estimate by 21% and 1%, respectively. The corresponding differences in the above ground biomass (AGB) estimates were 22% and 16%. However, only a few studies have used a sampling procedure with ALS data. Gautam et al. (2010) used a two-phase sampling procedure to estimate the forest AGB in Laos. The procedure integrates sample plots with ALS transects (10% coverage tested) and satellite images, and it attains a relative RMSE of 25 to 35 present in AGB in an area of 0.5 ha. A somewhat similar approach was used in Gregoire et al. (2011) and in Ståhl et al. (2011) for a large-area forest inventory in Norway. They used NFI field plots, ALS transects, and profiling LiDAR data. The two-phase laser sampling estimates for AGB were close to the estimates estimated using only field plots. In both studies, profiling LiDAR and ALS were also compared, and the ALS was found to be more useful.

The goal of this study was to investigate if the estimation accuracy could be improved through combining ALS data with optical satellite data. The tasks of the study were to choose an optimal number of strata and stratification.

2. Method

Ellipsoid stratification was used, which consists of two stages: the automatic derivation of stratum centroids and covariance matrices, and the assignment of observations. The population values are obtained as weighted averages of the sub-population values as

$$\hat{\bar{y}}_{STR} = \sum_{h=1}^{L} W_h \hat{\bar{y}}_h \qquad (1)$$

with an estimator of variance

$$var(\hat{\bar{y}}_{STR}) = \sum_{h=1}^{L} W_h^2 \, var(\hat{\bar{y}}_h) = \sum_{h=1}^{L} W_h^2 \frac{s_h^2}{n_h} - \sum_{h=1}^{L} W_h \frac{s_h^2}{N} \qquad (2)$$

where L is the number of strata, W_h is proportion of stratum h and s_h^2 is the sample variance within stratum h (Kangas et al., 2006):

$$s_h^2 = \sum_{i=1}^{n_h} \frac{\left(y_{h_i} - \bar{y}_h\right)^2}{n_h - 1} \qquad (3)$$

The numbers of strata studied were 5, 10, 15, 20, 25, 30, 35 and 40. In this study we utilized a Landsat 7 ETM+ satellite, and ALS from the Kuortane area in western Finland, covering an area of approximately 30 000 ha. The studied variable was the total stem volume for all species together. Sample plots, measured using a modification of the 10th National Forest Inventory measurement system, were used as field data. In our estimations, 6 Landsat 7 ETM image channels were used. In combining the two data sources, 6 Landsat channels, 23 laser scanning height-based extracted features from first echoes pulses and 4 intensity-based laser scanning features, e.g. Haralick texture features, were selected for the analysis.

2.1 Study area

The Kuortane study area is located in western Finland in the Southern Osthrobotnia region. It is covered mainly by middle aged boreal forest in the Suomenselkä watershed area. It is dominated by Scots pine (main tree species cover 81% of forest area), whereas Norway spruce and deciduous trees, mainly birches, usually occur as admixtures. The landscape is dominated

by pine forests growing on mineral soil and peatlands drained for forestry, open peatlands (mires) and agricultural fields in lower elevations. Terrain depressions of Kuortane are covered by lakes (largest Kuortanejärvi).

2.2 Data

The Finnish NFI is based on systematic cluster sampling design. Rectangular clusters of sample plots are located 7 km apart of each other in Central Finland, which includes the Kuortane region. Because the study area was only 20938.28 ha, the NFI sampling design was intensified for the present purpose. One cluster contains 18 field plots (temporary clusters) that are located along a rectangular tract, 300m apart. This gave a sample of 702 plots arranged in 39 clusters, of which 526 plots were located in the land use class forest. Each plot was measured as both a truncated angle count sample plot and fixed area plot, taking a basal area factor of 2 and a maximum radius of 12.52 m for the truncated angle count plots. The ALS data were collected on 28th July 2006 with an Optech 3100 laser scanning system operated at an altitude of 2 000m above ground level using a half-angle of 15º and a side overlap of about 20%. This resulted in a swath width of 1 070 m. The divergence of the laser beam (1 064 nm) was 0.3 mrad, which produced a footprint of 60 cm at ground level. Altogether 21 laser strips were measured, of which 2 were used for calibration purposes. The Optech 3100 laser scanning system produces four types of echoes (only, first, last, intermediate), which were re-classified as first and last pulse data. The Landsat 7 ETM+ scene was captured on 1st June 2007.

3. Results

The standard error for the total stem volume was found to decrease with an increasing number of strata. The ALS extracted features of the first pulse laser canopy heights notably improve the accuracy of estimation. Intensity-based Haralick texture features when used alone for stratification, performed reasonably well. The optimal number of strata is between 20 - 25. When all of the features from the two data sources are fused the results show an increase in uncertainty of the results due to the large number of features under investigation. A distribution of value of standard error depending on the number of strata is showed on the figure 1.

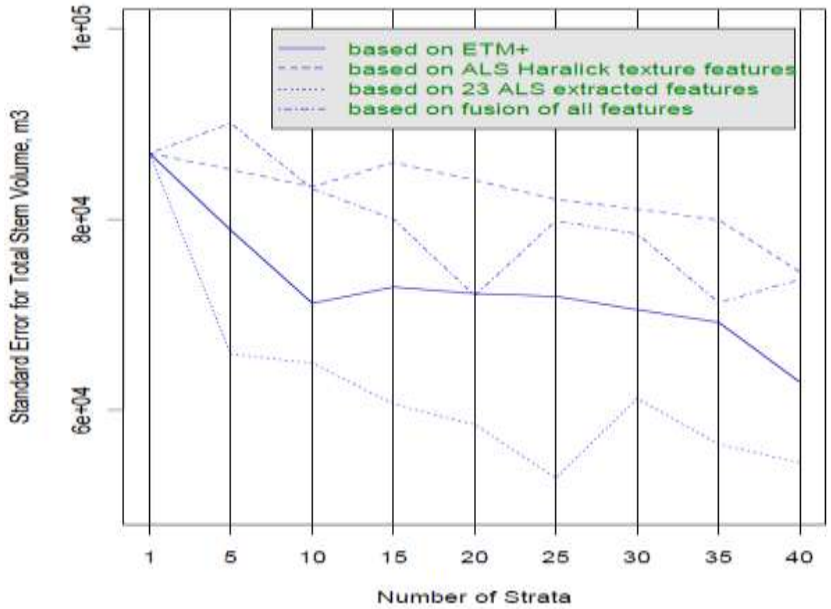


Figure 1: Standard error for the total stem volume for all species

Results of the study are presented on the table 1.

Table 1: Total stem volume and its standard error

number of strata	Based on ETM+ total stem volume, (m^3)/ %RSD	Based on ALS Haralick texture features total stem volume, (m^3)/ %RSD	Based on 23 ALS extracted features total stem volume, (m^3)/ %RSD	Based on fusion of all features total stem volume, (m^3)/ %RSD
1	1578521/ 5,51	1578521/ 5,51	1578521/ 5,51	1578521/ 5,51
5	1727411/ 4,57	1707346/ 4,99	1875970/ 3,51	1604161/ 5,61
10	1715378/ 4,15	1776744/ 4,69	1914989/ 3,39	1891467/ 4,40
15	1729210/ 4,22	1806269/ 4,76	1888907/ 3,21	1842848/ 4,34
20	1659859/ 4,35	1793756/ 4,69	1881926/ 3,11	1825147/ 3,94
25	1724103/ 4,17	1779898/ 4,61	1887130/ 2,80	1907555/ 4,19
30	1710880/ 4,12	1773397/ 4,57	1880386/ 3,26	1936080/ 4,05
35	1626122/ 4,26	1740087/ 4,59	1828692/ 3,09	1877640/ 3,80
40	1459334/ 4,31	1621177/ 4,60	1858626/ 2,93	1903362/ 3,87

4. Discussion and conclusions

In this study we evaluated the performance of ALS data in combination with satellite optical data for large-area forest stratification-based inventory in order to increase the estimation accuracy of total volume. We investigated the optimal number of strata and stratification variables for ALS-based stratification compared to satellite image-based.

We conclude that if ALS data is readily available it should be used in the stratification phase of the large-area forest inventory. The usefulness of fusing ALS data with optical data for stratification should be investigated further. The variable selection methods for the stratification based forest inventory should be tested.

Acknowledgements

This study was made possible by financial aid from Graduate School in Forest Sciences and the Finnish Academy project ‘Science and Technology Towards Precision Forestry’ (PreciseFor).

References

1. Gautam, B., Tokola, T., Hämäläinen, J., Gunia, M., Peuhkurinen, J., Parviainen, H., Leppanen, V., Kauranne, T., Havia, J., Norjamaki, I., & Sahet B., 2011. Integration of airborne LiDAR, satellite imagery, and field measurements using a two-phase sampling method for forest biomass estimation in tropical forests. International Symposium on “Benefiting from Earth Observation” 4 – 6 October 2010, Kathmandu, Nepal.
2. Gregoire, T., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R. & Holm S. 2011. Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research (41)*: 83–95
3. Holmgren, J. & Jonsson, T. 2004. Large Scale Airborne Laser Scanning of Forest Resources in Sweden. International Archives of Photogrammetry, *Remote Sensing and Spatial Information Sciences*, Vol. XXXVI – 8/W2. pp. 157–160.
4. Holopainen, M. & Hyyppä, J. 2003. Possibilities with laser scanning in practical forestry. In *Proceedings of the Scandlaser Scientific Workshop on Airborne Laser Scanning of*

Forests, ed. Hyyppä, J. Næsset, E, Olsson, H., Granqvist Pahlen, T. & Reese, H. 2003:264–273.

5. Kangas, A., & Maltamo, M., (2006). Forest Inventory - Methodology and Applications. Managing Forest Ecosystems 10. Springer. 364p. ISBN 1-4020-4379-1.
6. Nelson, R., Krabill, W. & Tonelli, J. 1988. Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment 24*: 247–267.
7. Nelson, R, Parker, G. & Hom, M. 2003a. A portable airborne laser system for forest inventory. Photogrammetric Engineering & Remote Sensing 69: 267–273.
8. Nelson, R Valenti, M.A., Short, A. & Keller, C. 2003b. A multiple resource inventory of Delaware using airborne laser data. *BioScience 53*: 981–992.
9. Nelson, R, Short, A. & Valenti, M. 2004. Measuring biomass and carbon in Delaware using an airborne profiling LIDAR. *Scandinavian Journal of Forest Research 19*: 500–511.
10. Nelson, R, Hyde, P., Johnson, P., Emessiene, B., Imhoff, M.L., Campbell, R. & Edwards, W. 2007. Investigating RaDAR-LiDAR synergy in a North Carolina pine forest, *Remote Sensing of Environment 110 (1)*: 98–108.
11. Næsset, E., 2004. Practical Large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research 19*, 164–179.
12. Næsset, E., , Gobakken, T. & Nelson, R., 2006. Sampling and Mapping Forest Volume and Biomass Using Airborne LIDARs. *Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium* 2006.
13. Ståhl, G., Holm, S., Gregoire, T.G., Gobakken, T., Næsset, E. & Nelson, R. 2011. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. *Canadian Journal of Forest Research 41*: 96–107.

The impact of lidar point density on classifying tree genus: using geometric features and vertical profile features

Connie Ko[1], Gunho Sohn[2] & Tarmo K. Remmel[3]

[1]Earth and Space Science and Engineering, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada, Tel: 416-736-2100 x20290, Fax: 416-736-5988; cko@yorku.ca

[2] GeoICT Lab; Earth and Space Science and Engineering Department, York University, Toronto, Canada, gsohn@yorku.ca

[3]Department of Geography, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada, Tel: 416-736-2100 x22496, Fax: 416-736-5988; remmelt@yorku.ca

Paper Number: ###SL2012-035

ABSTRACT:

We present a sensitivity analysis of the effect that LiDAR point density has on tree genus classification. Tree genus is classified by using two sets of features within the Random Forest classifier; the two sets of features are 1) descriptors related to the geometric information of the trees and 2) descriptors related to the vertical profile of LiDAR points. The targeted genera include pine, poplar and maple at a study site situated north of Thessalon, Ontario (east of Sault Ste. Marie) along a major utility corridor right of way and seven surrounding woodlots.

The first set of features classifies tree genera by obtaining the geometric information from LiDAR point data, we have derived 24 features for each isolated tree crown and the features are grouped into five categorizes 1) Linear features – describing characteristics of line segments derived internal to a tree crown; 2) Cluster features – describing the shape of the LiDAR point clusters grouped within an individual tree crown; 3) Hull features – describing tree form in terms of volume hulls; 4) Buffer features – describing characteristics of the point distribution within a tree crown and 5) Form features – describing overall tree shape. The second set of features classifies tree genera by obtaining vertical height point profile information, we derived 78 features for each tree; features are grouped in three categories. 1) Count features – percentage of first, single and last returns within the profile; 2) Height features – statistical attributes of height values at different height percentiles and 3) Intensity features – statistical attributes of intensity values at different height percentiles. The Random Forest classifier is used to process both sets of features separately and their accuracy rates are compared among different pulse densities. The LiDAR data was collected by a Riegl LMS-Q560 scanner with original point density of 40 pulses per m^2, tree genera classification was performed at the original pulse density, we then reduce the density to 20, 10, and 5 pulses per m^2, each time recording the accuracy. Our analysis indicates that as pulse densities decrease, vertical profile methods remain effective, but to truly extract structural parameters, higher pulse densities are necessary. The implementation of a simple multiple classifier system (MCS) combines the two sets of descriptors and yields improved classification accuracy. The sensitivity analysis helps identify a breakpoint at which pulse densities drop too low for true structural analyses at an operational scale.

1. Introduction

The objective of this paper is to analyse the impact of LiDAR (Light Detection and Ranging) pulse density on tree genus classification accuracy, using two types of classification features

jointly and severally. The first set of features is derived from the geometry of an isolated single tree point cloud and the second set of features is derived from vertical point profile metrics. Examples of using this type of metric for tree species/genera classification include Barilotti et al. (2009) fitting curved surfaces to canopy tops of different tree species; Kato et al. (2009) used an iso-surface algorithm to retrieve metrics regarding the shapes of tree crowns, volume, width, and base height. Vauhkonen et al. (2008, 2009, 2010) compute the alpha shapes of tree crowns as metrics for tree species identification. The benefit of using vertical profile metrics is that computation is relatively simple and therefore running the analysis is time effective; furthermore, rather than only using the location of the points (x, y, z values), it can take the intensity value into the consideration. Examples of using the vertical distribution (height and / or intensity) include Brandtberg (2007) and Kim et al. (2009) where the entire tree profile is considered. Additional consideration of attributes from height profiles is achieved by Ørka et al. (2009), Holmgren and Persson (2004), and Korpela et al. (2009). In our previous studies (Ko et al., 2012), where we compare the effectiveness of using geometric descriptors versus vertical profile descriptors for tree genera classification (*Pinus*, *Populus*, and *Acer*), we have shown that the classification accuracies by using the both sets of features are about the same (90% overall accuracy for both methods). The study was completed by using data that has 40 pulses per m^2 and therefore in this paper, we want to investigate the sensitivity of classification accuracy with the change in pulse densities. According to Vauhkonen et al. (2008) and Magnusson et al. (2007), the accuracies in estimating of tree attributes decreases when pulse density decrease. We assess the possibility of combining two methods and running our analyses with lower pulse densities and consequently lowering project costs.

2. Methods

2.1 Study area and LiDAR data

The study area is located about 75 km east of Sault Ste. Marie, Ontario, Canada. We chose eight field sites, one of the sites is an electric transmission line right-of-way (ROW) and we sampled the trees along the vegetation clearance zones. The rest of the field sites are chosen from the surrounding woodlots. In our study area, we have identified white birch (*Betula papyrifera* Marsh.), balsam fir (*Abies balsamea* (L.)), maple (*Acer saccharum* Marsh.), red oak (*Quercus rubra* L.), jack pine (*Pinus banksiana* Lamb.), poplar (*Populus tremuloides*), white pine (*Pinus strobus* L.), white spruce (*Picea glauca* (Moench Voss)) and others during field visits. Field validation was performed 30 July to 12 August 2009 and 8-10 August 2011, during the field visits, we validated 189 trees but only 160 of them belong to the genera of interest, therefore in this paper, we will perform classification analysis on only these 160 trees; our goal is to classify these trees correctly into their respective genera: pine, poplar and maple. Figure 1 shows an example tree with its vertical point distribution profile from each genus of interest, (a), (b), (c) are an examples of pine, poplar and maple respectively.

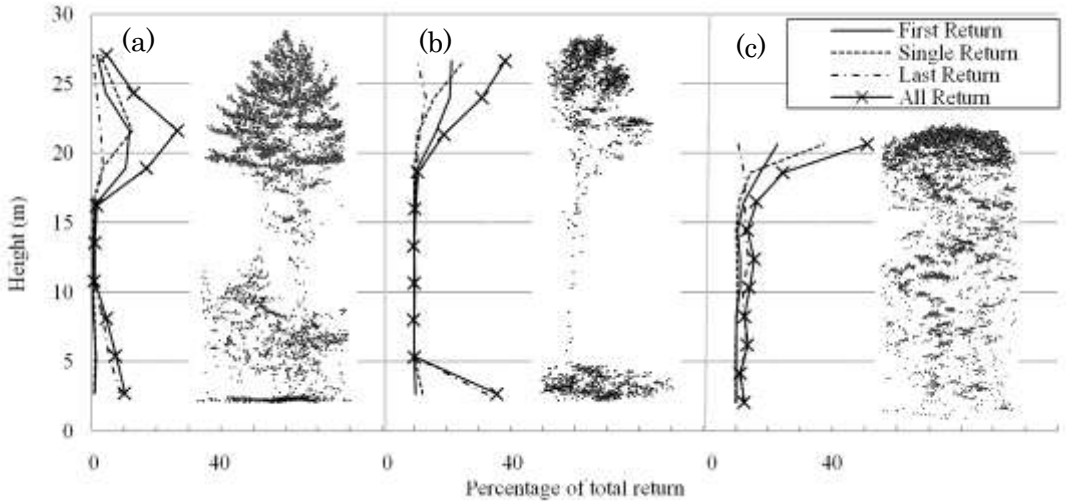


Figure 1: LiDAR data for pine (a), poplar (b) and maple (c) examples together with their corresponding vertical point distributions

LiDAR data was collected on 7 August, 2009 by a Riegl LMS-Q560 scanner, the flight altitude was approximately 122 m above ground level and the pulse density is approximately 40 pulses per m^2. The scanner has the capability of receiving up to five returns per pulse and each field site was covered with two flight paths, in opposing directions. That means each tree has up to 40 pulses × 5 returns per pulse × 2 flight paths = 400 points per m^2. Each tree is identified in the LiDAR scene by a GPS location obtained in the field and the trees are isolated from the mass point cloud by visual interpretation. The understory is removed by an algorithm discussed in Ko et al. (2010). Once the trees are isolated from the LiDAR scene with an original density of 40 pulses per m^2, they are thinned by removing every other point (with respect to gps time recorded by the scanner). Figure 2 shows an example of a pine tree with the various pulse densities.

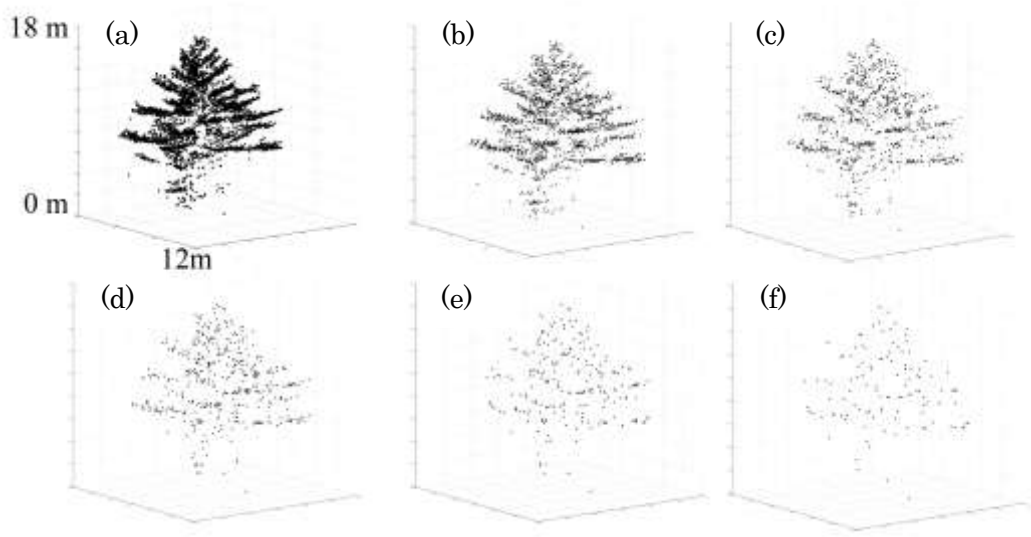


Figure 2: Example pine tree with (a) 40 pulses per m^2; (b) 20 pulses per m^2; (c) 10 pulses per m^2; (d) 5 pulses per m^2; (e) 2.5 pulses per m^2; (f) 1.25 pulses per m^2

2.2 Features descriptors and vertical profile descriptors for classification

We have derived two sets of features for classification; the first set involves grouping the LiDAR points within the tree crown into clusters; best fit lines and planes are then calculated for each cluster. The properties of these lines and planes are described in F1 to F10 (Table 1), properties related to the volume of the clusters are described in F11 to F17 (Table 1). 3D buffering features are features derived from expanding each LiDAR point in the tree crown into spheres by a distance of 2% of the observed tree crown height and are described in F18 to F21. F22 to F24 summarize the overall shape of the tree. Table 1 summarizes the qualitative description of the 24 features we have derived.

The second set of features summarizes the properties of vertical point distribution within each tree crown, we have included both height and intensity attributes. Table 2 summarizes the 78 vertical profile features.

Table 1: Descriptive summary of geometric derived features, source: Ko et al. (2012)

No.	Description
Line related	
F1	Average line segment lengths divided by tree height
F2	Average line segment lengths divided by tree crown height
F3	Average line segment lengths multiplied by the ratio between tree crown height and tree height
F4	Average line segment angles measured from the x-y plane to the line
F5	Average line segment angles measured from the y-axis to the line projected onto the x-y plane
Cluster related	
F6	Average number of points in each cluster divided by the number of points in the tree crown
F7	Average of the average orthogonal distance from each point to the line in the tree crown for each cluster
F8	Average of the average orthogonal distance from each point to the plane in the tree crown for each cluster
F9	F7 divided by the tree crown height multiplied by F8 divided by the tree crown height
F10	Average of the volume of the convex hull for each cluster divided by the number of points in the cluster
Convex hull and alpha shape related	
F11	Average of volume of the convex hull for each cluster divided by the number of points in the cluster
F12	The difference between the area of the convex hull and the alpha shape compared to the convex hull area
F13	Volume of the tree crown convex hull divided by the number of points in the crown
F14	Volume of the tree crown alpha shape divide number of points in the crown
F15	Average distance from each point to the closest facet of the convex hull
F16	Standard deviation of orthogonal distances from each point to the convex hull
F17	Coefficient of variation (F15/F16)
3D buffering related	
F18	Sum of overlapped volume between i^{th} and j^{th} spheres
F19	Overlapped count of points captured by i^{th} and j^{th} spheres
F20	Overlapped volume divided by the number of points in the tree crown
F21	Number of count divided by the number of point in the tree crown, squared
Overall tree shape related	
F22	Tree height divided by the radius of the tree crown, radius is obtained by assuming when the tree crown is projected to xy plane, it is circular
F23	Tree crown height divided by the radius of the tree crown, radius is obtained by assuming when the tree crown is projected to xy plane, it is circular
F24	Tree crown height divided by the tree height

Table 2: Summary of features derived from vertical point profile: a= first; s= single; l= last; Std= standard deviation; Cv = coefficient of variation, source: Ko et al. (2012)

10^{th} percentile	**50^{th} Percentile**	**90^{th} percentile**
% of canopy return ($V1_f$, $V2_s$, $V3_l$)		
% return count ($V4_f$, $V5_s$, $V6_l$)	% return count ($V7_f$, $V8_s$, $V9_l$)	% return count($V10_f$, $V11_s$, $V12_l$)
Mean height ($V13_f$, $V14_s$, $V15_l$)	Mean height ($V16_f$, $V17_s$, $V18_l$)	Mean height ($V19_f$, $V20_s$, $V21_l$)
Mean height of canopy return ($V22_f$, $V23_s$, $V24_l$)		
Std of height ($V25_f$, $V26_s$, $V27_l$)	Std of height ($V28_f$, $V29_s$, $V30_l$)	Std of height ($V31_f$, $V32_s$, $V33_l$)
Std height for canopy return ($V34_f$, $V35_s$, $V36_l$)		
Cv height for canopy return ($V37_f$, $V38_s$, $V39_l$)		
Kurtosis of variation height for canopy return ($V40_f$, $V41_s$, $V42_l$)		
Skewness of variation height for canopy return ($V43_f$, $V44_s$, $V45_l$)		
Mean intensity ($V46_f$, $V47_s$, $V48_l$)	Mean intensity ($V49_f$, $V50_s$, $V51_l$)	Mean intensity ($V52_f$, $V53_s$, $V54_l$)
Mean intensity of canopy return ($V55_f$, $V56_s$, $V57_l$)		
Std of intensity ($V58_f$, $V59_s$, $V60_l$)	Std of intensity ($V61_f$, $V62_s$, $V63_l$)	Std of intensity ($V64_f$, $V65_s$, $V66_l$)
Std intensity of canopy return ($V67_f$, $V68_s$, $V69_l$)		
Cv intensity of canopy return ($V70_f$, $V71_s$, $V72_l$)		
Kurtosis of variation intensity of canopy return ($V73_f$, $V74_s$, $V75_l$)		
Skewness of variation intensity of canopy return ($V76_f$, $V77_s$, $V78_l$)		

2.3 Classification by Random Forest

The Random Forest classifier is an ensemble classifier that combines numerous weak classifiers, in this case, classification trees to become a strong classifier (Breiman, 2001). ⅓ of the original data (out of bag sample) is left out of the process for the validation procedure. The number of features selected to construct each tree is a user defined number and by replacing the classification tree with different features, it will result a different validation error (out of bag error), and the change of error determines whether a particular feature improves or degrades the overall classification error and the change is called the mean decrease accuracy. This is recorded to estimate the importance of each feature.

Unlike the traditional method of using Random Forest where ⅔ of the data are split into in-bag samples (data for training the classification trees) and ⅓ of the data are out-of-bag samples (data for validation), we split our data into 5% increment partitions for training and validating. We start by using 5% of the data for training (95% of the data for validation) through to 95% of the data for training and run Random Forest 20 times at each partitioning by re-selecting the training and validation tree samples to obtain the average classification accuracy. The average mean decrease accuracies for all the features are also obtained to have a better understanding of the importance of the features. The goal of this experiment is to determine the best partition for training; that is to minimize the amount of data for training and yet produce reasonable classification accuracies.

To study the potential of collaborating two classifiers (two sets of features) to achieve higher classification accuracy, we combined the two classifiers by using vertical profile features as a base classifier, followed by using geometric features as supplementary information for making

the final classification decision. First, we use 25% of the data for training and predict the 75% of trees by vertical profile features with Random Forest, this process is repeated with geometric features. Then, the margins for each tree are calculated and normally the class with the largest margin represents the prediction with the most confidence and the tree will be assigned to that class. Instead, we calculate the difference between the largest margin and the second largest margin from the vertical profile features classifier and we select the trees with the difference less than 0.5 because small difference in margin between the first choice and the second choice indicate trees that are potentially mis-classified and these are the tress that needs to be reconsidered by geometric features. Next, within the subset of trees, we separate the trees into two groups; trees that has same decisions made by vertical profile and geometric features and trees that has different decisions. Trees that have synchronized decisions will be kept as is and trees that have different decisions will be replaced by the decision made by the set of features that has a larger difference between the first and second margins. New overall classification accuracy is updated.

3. Results

The graphs of classification accuracies with different partition levels is displayed in Figure 3, Figure 3(a) shows the classification accuracies with different pulse densities by using the geometric descriptors and (b) shows the classification accuracies with different pulse densities by using vertical profile descriptors, each partition of the graph below represent an average classification accuracy of running Random Forest 20 times.

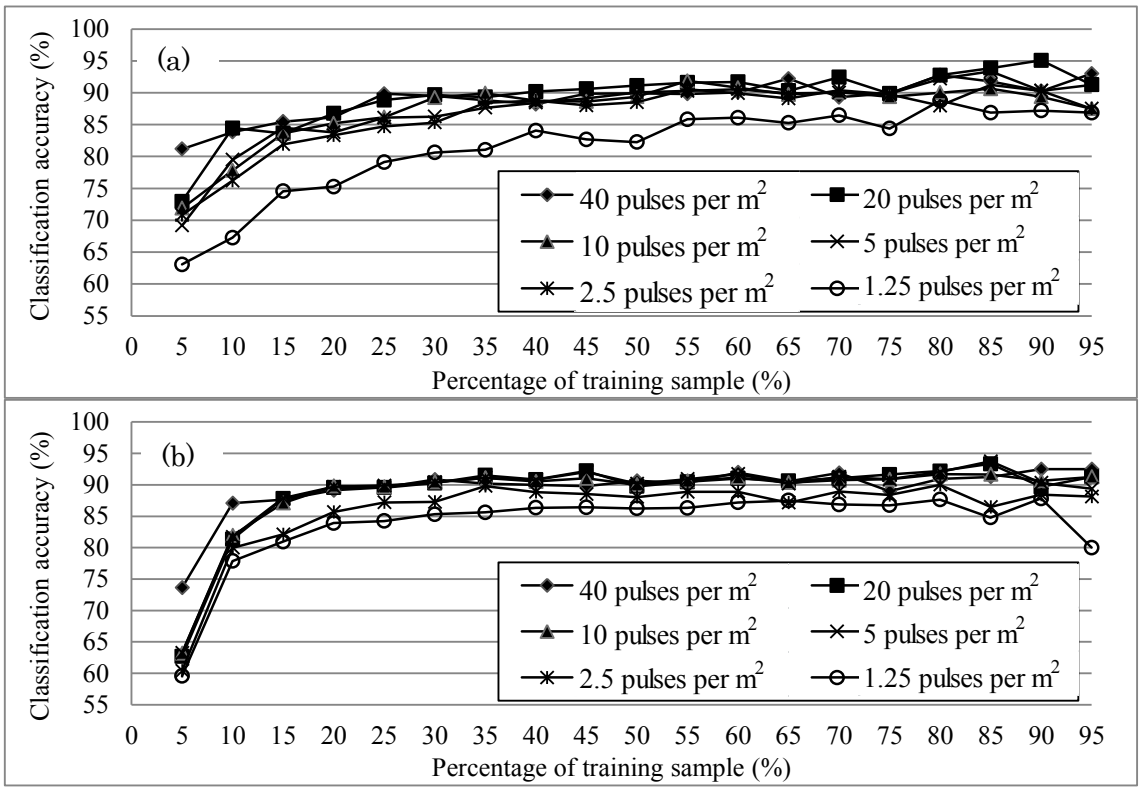


Figure 3: (a) The average change in the classification accuracy with different training sample partition using geometric descriptors (b) Average change in classification accuracy with different pulse densities by using vertical profile descriptors

The classification accuracy increases for both sets of descriptors with various pulse densities as

the amount of data partitioned for training increases. In this paper, we chose the partition level using 25% of the data for training and 75% of the data for validation, at this partition, both methods have similar classification accuracies (both round to 90%) at the original density level (40 pulses per m^2).

Figure 4 shows the relationship between classification accuracy with different pulse densities for geometric descriptors, vertical profile descriptors as well as the combined results. The values for the accuracies are included rounded to the nearest hundredth.

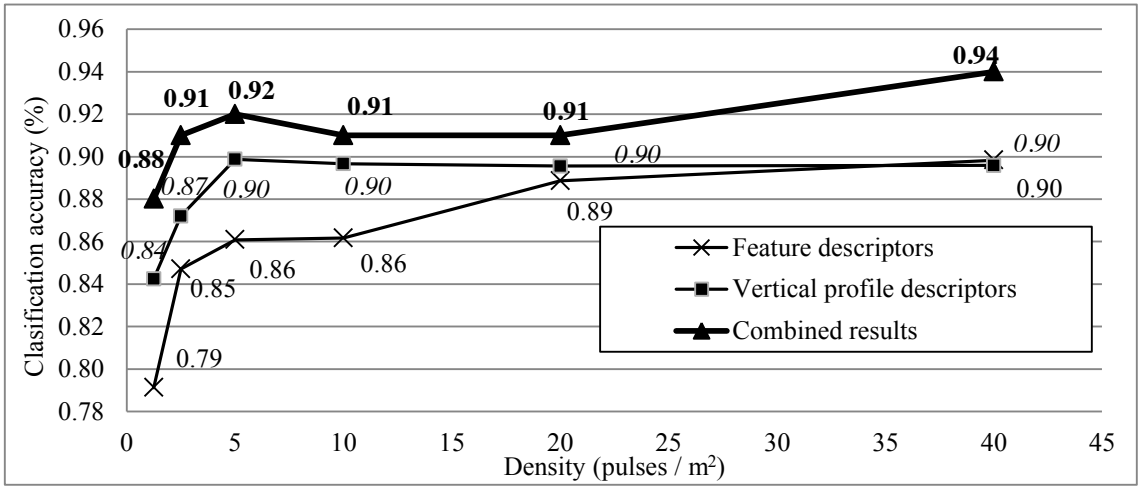


Figure 4: Classification accuracy with different LiDAR pulse densities for geometric descriptors, vertical profile descriptors and the combined results

By calculating the mean decrease accuracy, we can assess the importance of each feature in relation to the other. The two most important features for geometric descriptors are F24 and F3 (ratio between tree height and crown height and average line segment lengths in relation to crown height and tree height ratio, respectively) for 1.25, 2.5, 5 and 10 pulses per m^2; F24 and F20 (average overlapped volume by 3D buffering per point) for 20 and 40 pulses per m^2. The two most important feature for vertical profile descriptors is V10 and V45 (percentage of first return at 90^{th} percentile, skewness of height variation for canopy last returns) for 1.25, 2.5, 5, 10, 20 pulses per m^2 and V11 and V44 (percentage of single return at 90^{th} percentile, skewness of height variation for canopy single returns) for 40 pulses per m^2.

4. Discussion

Figure 3 shows that the learning rates generally increase as the available data for training increases. By using only using 25% of the data for training, both methods achieved 90% accuracy rate implied that the default Random Forest partition is excessive in our example. By comparing both descriptors, when there are less than 20% of the data available for training, vertical profile descriptors provide better classification accuracy. Moreover, in Figure 3(a), there is a relative bigger gap between 1.25 pulses per m^2 with the rest of the densities, this shows the geometric descriptors has a drop in efficiency at that density level.

We have selected the 25% partition level to examine the change in accuracy relative to changes in pulse density. Figure 4 shows that classification accuracy generally increases when density increases as there is more information available for processing. The increase slows down for all methods meaning there is room for reducing the pulse density to trade for slightly lower classification accuracy, and normally results in lower costs. Geometric features are more

complex and therefore requiring a higher point density. This relationship is being shown in Figure 4 where geometric descriptors achieve 90% accuracy at a higher density level. However, geometric features are able to correct some of the mistake that is being made by vertical profile descriptors, the combined classifiers shows the highest accuracy in Figure 4.

The important features among geometric descriptors are F24, F3 and F20. For F24, this is because poplar tends to have the smallest ratio (relatively small tree crown size) and maple tends to have the largest. F3 characterizes the length of the line segments in relation to the normalized tree crown size, lines derived for poplar tends to be short, lines for pines tends to be longer and representative of the branches. Maple has the longest derived lines. F20 highlights the spatial distribution of points within the tree crown. Pines and maples have high values because for pine, points tends to cluster at the layers of branches and for maple, points tends to cluster at the top of the tree crown, leaving poplar having the lowest value because the spatial distributions of points inside poplar tree crowns are even. For vertical profile descriptors, V10 and V11 are important and this coincide with the result obtained from Holmgren and Persson (2004), where the important features are located on the upper percentiles of the tree. V44 and V45 refers to the skewness of height distribution within the entire tree canopy (single and last, respectively), the canopy for maple trees are closer, resulting a higher ratio of single returns on the top of the trees therefore the distribution of points are more skewed. Pine and poplar trees allow LiDAR pulses to reach closer to the ground level and therefore more of the last returns are closer to the ground level.

This paper compares the performance of classifying tree into three genera by using two different approaches separately and then an attempt to combine the two methods. The advantages of using geometric descriptors are that its close relationship of the features to the biophysical interpretation of trees, it also yields a sensational visual interpretation. We also observed that most of the errors from vertical profile descriptors comes from poplar – pine classification, this is because sometimes, the point distribution for these two genera are similar but geometric descriptors yield better accuracy at this situation. By combining the decisions made by the two classifiers can improve classification accuracy, the results shows that some of the mistakes can be corrected automatically by implementing a simple Multiple Classifier Systems (MSC). Further and on-going research is studying the different ways of fusing the classifiers effectively in order to improve classification accuracy.

Acknowledgements

This research was funded by GeoDigital International Inc., Ontario Centre for Excellence (OCE) and NSERC.

References

Barilotti, A., F. Crosilla, and F. Sepic, 2009. Curvature analysis of LiDAR data for single tree species classification in alpine latitude forests. *Laser scanning*, 1-2 September 2009, Paris, France (IAPRS, Vol. XXXVIII, Part 3/W8).

Brandtberg, T., 2007. Classifying individual tree species under leaf-off and leaf-on conditions using airborne LIDAR. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(5), 325–340.

Breiman, L., 2001. Random forests. *Machine Learning*, 45(1), 5-32.

Holmgren, J., and Å Persson, 2004. Identifying species of individual trees using airborne laser scanning. *Remote Sensing of Environment*, 90(4), 415-423.

Kato, A., L.M. Moskal, P. Schiess, M.E. Swanson, D. Calhoun, and W. Stuetzle, 2009. Capturing tree crown formation through implicit surface reconstruction using airborne lidar data. *Remote Sensing of Environment*, 113(6),1148-1162.

Kim, S., R.J. Mcgaughey, H. Andersen, and G. Schreuder, 2009. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment*, 113(8),1575-1586.

Ko, C., G. Sohn, and T.K. Remmel, 2012. A comparative study using geometric and vertical profile features derived from airborne LiDAR for classifying tree genera. *The XXII Congress of the International Society of Photogrammetry and Remote Sensing*, 25 August - September 1, Melbourne, Australia.

Ko. C, T.K. Remmel and G. Sohn, 2010. A statistical partitioning of vegetated airborne laser scanning data towards understory and canopy separation. *Proceedings of the Canadian Association of Geographers Annual General Meeting*. June 1-5, Regina, Saskatchewan, Canada, 177-180.

Korpela, I., T. Tokola, H.O. Ørka, and M. Koskinen, 2009. Small footprint discrete-return LiDAR in tree species recognition. *Proceedings of the ISPRS*, 2-5 June 2009, Hannover, Germany (Workshop).

Magnusson, M., Fransson, J. E.S. and J. Holmgren, 2007. Effects on estimation accuracy of forest variables using different pulse density of laser data. *Forest Science*, 53(6), 619-626.

Ørka, H.O., E. Næsset, and O.M. Bollandsås, 2009. Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment*, 113(6), 1163-1174.

Vauhkonen, J., I. Korpela, M. Maltamo, and T. Tokola, 2010. Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sensing of Environment*, 114(6), 1263-1276.

Vauhkonen, J., T. Tokola, M. Maltamo, and P. Packalén, 2008. Effects of pulse density on predicting characteristics of individual trees of Scandinavian commercial species using alpha shape metrics based on ALS data. *Canadian Journal of Remote Sensing*, 34(2), S441-S459.

Vauhkonen, J., T. Tokola, P. Packalén, and M. Maltamo, 2009. Identification of Scandinavian commercial species of individual trees from airborne laser scanning data using alpha shape metrics. *Forest Science*, 55(1), 37-47.

Introducing Info Clouds as an alternative to LiDAR

Robert Uebbing

North West Group, robert.uebbing@nwgeo.com

Paper Number: SL2012-036

1. Introduction

In the last decade airborne LiDAR (Light Detection and Ranging) has established itself as a key technology to capture high resolution Digital Surface Models (DSM) and Digital Elevation Models (DEM). Over the same time the demand for higher accuracy and higher resolution DSMs and DEMs, delivered more frequently, has increased. LiDAR is a natural choice to fulfil this demand for various applications. However, the high acquisition costs often prohibit its use for larger areas of interest. Considering the fact that imaging sensors like Leica Geosystems' ADS provide multiple stereo coverage, the image data itself can and should be used for photogrammetric DSM derivation. Therefore, in addition to LiDAR data acquisition and processing, North West Geomatics has developed a DSM generation tool from ADS line scanner imagery in cooperation with Leica Geosystems.
Generating a very high DSM andDEM resolution, in the order of the image ground sampling distance (GSD), presumes matching at the actual image resolution; or in other words: utilizing a per-pixel matching cost. Depending on image texture, a per-pixel measure is generally ambiguous; additional constraints, such as the assumption of a smooth surface, need to be introduced. Algorithms that globally minimize both cost and constraints are called global image matching; they are among the top-ranked matching approaches in terms of quality and resolution. Their drawback is performance, which is addressed by the Semi-Global Matching (SGM) approach of Hirschmüller (2005, 2008). SGM approximates the two-dimensional, global aggregation of matching cost by a number of one-dimensional cost paths. It still achieves similar accuracy as truly global matching but it is significantly faster – refer, e.g., to the systematic comparisons of SGM with local and global matching algorithms using different cost functions by Hirschmüller and Scharstein (2007, 2009). As a result, SGM has been further investigated and also enhanced by various researchers for different applications and data sets, including aerial images (Hirschmüller et al., 2005, Hirschmüller, 2008), terrestrial and extraterrestrial satellite data (Hirschmüller et al., 2006, Krauß et al., 2008, Alobeid et al., 2009) or video sequences (Heinrichs et al., 2007, Gerke 2008, Hermann et al., 2009). Furthermore, SGM is being recognized and deployed in the industry. The SGM approach fulfils our need of high DSM resolution and high performance allowing the generation of surface models for larger regions such as states or provinces. Based on the quickly gaining mindshare and promising SGM results, we adapted the algorithm to the unique properties of ADS line scanner imagery using our existing software framework, in particular the highly optimized ADS sensor model. The remainder of this paper describes the DSM derivation from ADS data based on SGM. The properties of SGM DSMs and derived DEMs are compared to LiDAR products, both in general and for an example data set that has been captured over the municipality of Romanshorn, Switzerland. Also, the "Info Cloud" as a fundamental data container for downsteam applications and work flows is introduced in this context.

2. Semi-Global Matching for ADS

SGM is a new image matching approach, which originates from the computer vision community. It has been developed by Hirschmüller (2005, 2008). The core algorithm aggregates per-pixel matching costs under consideration of smoothness constraints. The minimum aggregated cost leads to the disparity map (or, respectively, parallax map) for an image stereo pair, providing the corresponding location for each base image pixel on the respective epipolar line in the pair image. The SGM approach is suited for DSM collection in very high resolution, i.e. the image GSD.

2.1 Matching Cost: Difference of Intensity Gradients

The difference of intensities of allegedly corresponding pixels (in global matching including SGM) or search windows (in local methods) is a simple but widely used cost function: The smaller the difference, the more likely is the correct match – provided that overall intensity offsets in the images of a stereo pair are removed. One solution is the Sobel filtering of the imagery as proposed and tested by Hirschmüller & Gehring (2009) with different matching approaches including SGM. The Sobel filter is applied solely in the direction of the epipolar lines, which corresponds to the flight direction or x in the ADS image coordinate system. This convolution has two effects: First, radiometric offsets are removed. Second, the Sx eliminates y gradients or, respectively, horizontal structures that might have a negative impact in matching in case of (unavoidable but small) errors in the image orientation. However, our main driver for switching the cost from Mutual Information (MI) that we initially used for ADS processing (Gehrke et al., 2010) has been occasional errors; in brief: MI cost reflects the probability of a certain pair of intensities to be a valid match, and for a particular project in North West Geomatics' production the probability computation was dominated by grassland and, applied to the neighboring road (with similar intensities but slightly different pair matches due to a different BRDF), elongated outliers occurred despite the SGM smoothness constraints. Already used for large projects of various type – including all examples presented in this paper –, the Sobel-based cost has not yet shown any issues on flat structures and open areas in general. This is crucial when aiming for the derivation of a DEM.

2.2 Cost Aggregation and Disparity Computation

Any pixel-wise matching cost measure is generally ambiguous. Therefore, constraints such as the assumption of a smooth surface need to be introduced and globally aggregated along with the cost. The Sobel-based cost c is computed for all potential matches in a stereo pair, i.e., for each base image pixel p and for all disparities d that provide the link to pair image locations. This results in cost vectors for each pixel and, in a cost cuboid for the entire image. The energy function E(D) that has to be globally minimized consists of this cost and constraints on smoothness. Hirschmüller (2005, 2008) proposes to use the penalties P_1 for slight variations of one disparity and P_2 for any larger changes, i.e. discontinuities, in-between neighboring pixels:

$$E(D) = \sum_p e(p,d) = \sum_p c(p,d) + \sum_{q \in N_p} P_1 T(|d - d_q| = 1) + \sum_{q \in N_p} P_2 T(|d - d_q| = 1) \quad (1)$$

The T(…) function controls the application of penalties, i.e. the weighting between matching and smoothness constraints; it is 1 in case that its argument is true and 0 if false. SGM approximates the theoretically desired two-dimensional, global aggregation of matching cost by a number of one-dimensional cost paths for each potential disparity. A total of eight paths at every 45° is considered sufficient and also used for ADS processing (Gehrke et al., 2010). The overall cost vector for each base image pixel results from the summation of these eight aggregated cost vectors at that pixel. Based on that, the final disparity map is provided by the disparities corresponding to the minimum of each pixel's summed cost vector. It is generated for the base image and also for the pair image by switching the role of base and pair images. This reverse matching allows for the elimination of mismatches, which occur mainly in occluded areas.

2.3 ADS Line-Scanner Data: Considerations and Benefits

The ADS captures multiple panchromatic stereo image bands, which provide redundancy to fill gaps from occlusions as well as the ability for consistency checks between stereo pairs. Full colour (RGB) as well as NIR bands provide further stereo coverage along with additional radiometric properties, which can be beneficial for matching and also support post-processing (classification) of the results. The unique sensor geometry has to be regarded.

2.3.1 Epipolar Geometry

The projection of line scanners is central within each line and parallel in flight direction, with every scan-line featuring its individual exterior orientation. Even though a 'perfect', linear flight would result in straight, parallel epipolar lines, aircraft movements distort the imagery and accordingly the epipolar geometry. Most of these distortions are eliminated by pre-rectifying the original image data (L0) onto a plane (L1). While SGM is expected to generally benefit from using such L1 data, the major advantage over L0 is the possible approximation of the resulting epipolar curves as piecewise linear features. Similarly, these epipolar curves can be computed in a sparse pattern (4-16 pixels) with linear interpolation for pixels in-between, which makes both the epipolar grid initialization – based on the highly optimized ADS sensor model – and the time-critical epipolar point computation from disparity very fast.

2.3.2 Object Point Computation

A side-effect of the described epipolar grid setup is the intermediate step of 3D point computation for each line segment. Based on those points, final object points for all base image pixels can be derived by linear interpolation depending on disparity. This is significantly faster than performing forward intersections.

2.3.3 Multiple Stereo Coverage

Most configurations of the ADS provide three panchromatic stereo angles, in case of the ADS80: 14° backward, 2° forward and 27° forward. This results in three possible stereo pairs, although we do not use the largest viewing angle combination. That leaves two stereo pairs, with the nadir-most viewing angle providing the common base image. Then, either object points or disparity images can

be joined, the latter of which requiring an intermediate object point computation. Although this seems extra work, the 'detour' of joining disparity images is preferred since it eases subsequent steps such as outlier elimination.
In this context it should be mentioned that the matching cost permits the use of the ADS colour bands, which increases stereo coverage and potential matching quality (based on different radiometry). In rare circumstances such as CCD malfunction or oversaturated panchromatic bands colour bands have been utilized for production work.

2.3.4 Handling of Large Image Takes

ADS L1 images are more than 12,000 pixels wide and can be up to 1,000,000 pixels long. This means that SGM processing needs to be carried out on sub-tiles of, e.g., 1,024 x 1,024 pixels. Considering the ADS viewing angles and resulting stereo pairs, this size is just small enough to model a mountainous area or a downtown core in a fairly high GSD of 0.10 m – which might require in the order of 600-800 disparities – on a 64-bit system with 8 GByte memory; see Hirschmüller (2005) for details of SGM memory usage. Based on the fact that results from a large number of takes have shown no evidence of tiling effects, the sub-tiles are processed independently without any overlap (apart from the area that is clipped already during processing – see above).

3. Comparison between LiDAR and SGM

LiDAR-acquired DSMs have become the standard for high-resolution surface generation. As an active sensor, LiDAR data can be captured regardless of light conditions and the laser pulses can penetrate into forest canopy to measure the ground directly – a distinct advantage over competing techniques such as the image-based SGM. In addition to the elevation, LiDAR points often contain intensity information from the laser pulse, which can be plotted as an orthoimage. Due to a combination of sensing methods and geometry, LiDAR has a greater accuracy in height than in horizontal position; for high precision work typically 5 cm in Z and 10-15 cm in XY (Morin and El-Sheimy, 2001a). LiDAR point accuracy can also be affected by unmodeled atmospheric conditions (such as mist or volcanic ash) or by the reflectance properties of the target (Morin and El-Sheimy, 2001b).
As an emerging technology, SGM-derived DSMs are intended to be used as an alternative source of elevation and as a premier choice for ultra-dense surface extraction. DSMs are computed from imagery, which is typically acquired under strict mission constraints (sun angle, cloud cover, etc.); but as a derived product it does not add any cost to an imaging flight. In contrast to LiDAR, which allows storing the LiDAR intensity of each terrain point, SGM points can be assigned the full range of spectral bands, typically the red, green, blue and near infrared, based on the ADS source imagery. They can be plotted as an orthoimage. SGM accuracies are driven by the triangulation accuracies of the imagery – typically 0.5 GSD horizontally and 1.5 GSD vertically. In the case of ADS line scanner imagery, extreme turbulence can reduce the quality of the images and hence the derived SGM DSM.
For providing a DSM in a certain resolution, the acquisition time and cost for LiDAR is generally higher than for imagery, which forms the basis for SGM. Processing performance can vary widely for LiDAR and SGM, but in general SGM has a larger computation cost than LiDAR. This can be mitigated through the use of multi-threading, large scale cluster deployment and graphic processing units (GPUs) to accelerate the calculations. With the current CPU implementation, the SGM processing achieves 10,000-20,000 points/s (Intel Core i7 @ 2.8 GHz), depending on disparity range. LiDAR processing speed is in the order of 1,000,000 points/s (2 x AMD Opteron 2220 @

2.80 GHz; Leica Geosystems, 2009). Data can be classified automatically to determine ground (DEM), buildings or vegetation content with minor changes to classification rules to accommodate the different point densities. The potentially higher density of the SGM DSMs can ease identification of structures in the data, making manual editing less error prone. On the other hand, LiDAR's ability to penetrate canopy and supply multiple returns enhances its ability to determine bare earth.
For the purpose of orthoimage generation, in addition to the obvious advantage of saving data acquisition costs there are distinct benefits of using image-based DSM generation: The geometry (datum) of the DSM is identical to the geometry of the images. Any adjustments made to the imagery – such as the triangulation – will be reflected in the DSM. Also, as the imagery and the DSM reflect the same time of recording, which means that the representation of the earth's surface is consistent; any inaccuracies and noise due to temporal changes can be avoided. Lastly, given that the DSM is directly derived from the imagery, the radiometric properties can be assigned to each generated 3D point, thus creating an "Info Cloud" that contains the geometry and the radiometry aligned perfectly in the same dataset.

4. Evaluation Results

SGM evaluation was undertaken using data from the Romanshorn municipality at Lake Constance in northern Switzerland. The imagery has been captured in the spring of 2007 using a Leica Geosystems ADS40, acquiring images at 5 cm ground sample distance (GSD). The LiDAR DSM has been captured in the autumn of 2007 using a Leica Geosystems ALS50-II at a nominal point spacing of 30 cm. Imagery was adjusted with photogrammetric control and constrained with inertial navigation data (Sun et al., 2006). The estimated pixel accuracy is 2.5 cm in XY and 7.5 cm in Z. LiDAR data was adjusted to remove systematic errors (Morin and El-Sheimy, 2002) and co-registered with the imagery. The estimated LiDAR point accuracy is 15 cm in XY and 5 cm in Z. A subset of the area was chosen for SGM testing. It consists of a region with mixed urban areas, fields and water, 375 m x 475 m in size (Figure 1). There were ~1.5 million LiDAR points in this area. SGM resulted in ~63 million points based on ~74 million image pixels, which translates to 85% yield. To compare surfaces, the SGM and LiDAR DSMs were analysed in a commercial DSM package. Elevation differences were generated from unedited and ground-classified models and plotted to observe any systematic effects. Profiles were taken to highlight differences in urban areas (buildings, roads, etc.) and difference sta-
tistics calculated to judge performance. The SGM-LiDAR comparison shows an average difference of 0.8 cm ± 5.4 cm. The small standard deviation suggests that the imagery and/or LiDAR data are more accurate than expected and that SGM potentially maximized its vertical accuracy. The difference plot in Figure 2 shows a number of areas (coloured white) larger than 3 GSD. In most cases these represent temporal effects such as grass pasture seasonal changes (large areas to lower left), construction (large rectangles) or changes in trees (random roundish shapes). Black areas contain no SGM data, mainly caused by occlusions; they naturally occur next to buildings and trees and, due to the sensor geometry, increase towards the edges of the flight line, which are top and bottom in the difference plot.

Figure 1. Romanshorn Test area RGB

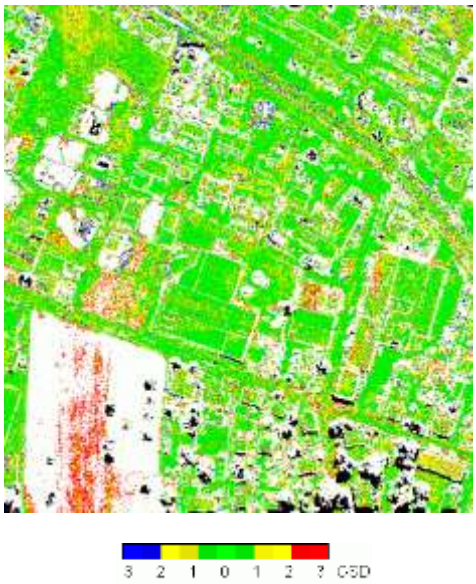


Figure 2. Difference between SGM and LiDAR

The strong agreement between LiDAR and SGM results is supported by the profile (Figure 3), which shows a very good agreement between LiDAR and SGM points along man-made structures, and demonstrates some of the differences in trees while canopy heights are generally similar.

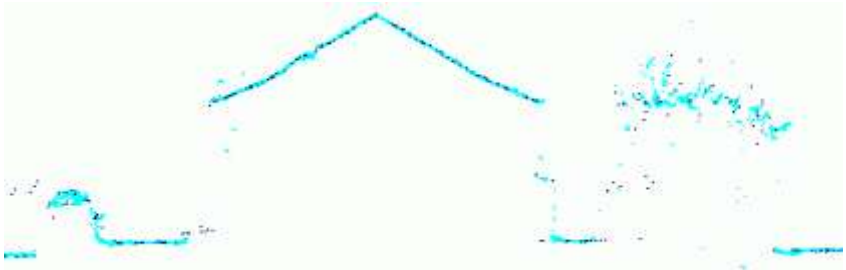
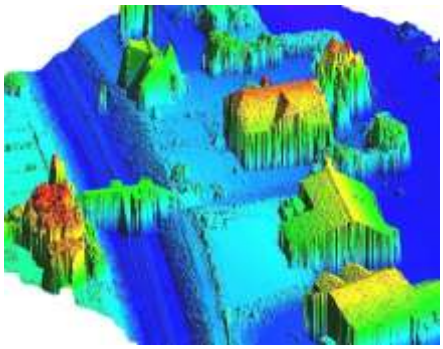


Figure 3. Profiles through SGM results (light blue) and LiDAR (dark blue)



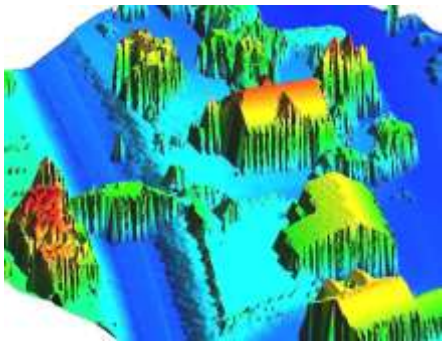


Figure 4. SGM (perspective view) Figure 5. LiDAR (perspective view)

A strength of pixel-based matching approaches like SGM is the behaviour at discontinuities, which is especially important in urban areas. In the comparison of perspective views – Figure 4: SGM, Figure 5: LiDAR – it can be seen that SGM has maintained sharp edges and also resolved a lot of detail on structures such as the railway tracks or the post next to them, which is not contained in the LiDAR data. (The object next to the railway in the LiDAR view – consisting of a single point – but not contained in the SGM result is a bush. It indicates the LiDAR advantages at vegetation, whereas SGM might have not properly matched it and/or classified it as an error.)

5. LAS Output: The Info Cloud

The results of SGM including the automatic post-processing are projected into 3D space in a user-definable coordinate system. At the time of this writing, they are stored in the LAS 1.2 format as defined by the ASPRS (2008). Each DSM job will generate its own output, one file for the full resolution data that consists of 10-20 million points and an additional file for the thinned data that might still include up to about 2 million points, depending on initial coverage and thinning rate. Results can be generated in point data record formats 1 or 2. However, only some of the foreseen items, which are designed for LiDAR, are assigned from ADS/SGM results. In point data record format 1, each object is assigned its coordinates, the intensity of the corresponding base image pixel and a default classification (ground). We also provide the ability to store ADS color data. Those are generated in an on-the-fly rectification by projecting each object points into the red, green, blue, and near infrared color bands and assigning the interpolated, radiometric corrected intensities. Since the LAS version 1.2 only allows storing three color components, two LAS files are generated – the first file is encoded with RGB and the second file is encoded with FCIR. It is to be noted that the new LAS 1.4 specification allows 4 colour components, thus eliminating this constraint. Both files contain the identical geometric point cloud, which leads to altogether four files: High resolution RGB and FCIR and thinned RGB and FCIR. With this color information and the panchromatic intensity, the high density point cloud becomes a new type of product: the multi-band ‘Info Cloud’. Examples of which are shown in Figure 6.

Figure 6. Examples of Info Clouds displayed in 3D perspective

6. Info Cloud: Usage in the forestry industry

Automatic tree counting and type specification has been a challenge in the forestry industry. At Silvacom Group an approach that combines LiDAR with ortho-photos has been in use. The approach uses classification methods to determine forest stands and their type in the ortho-photos and correlates their location with the 3D point cloud derived from LiDAR. This approach has 2 major challenges that are eliminated by using Info Clouds from SGM. The first issue is the co-registration between the LiDAR and the ortho-photos which remains difficult (Kato, et al. 2010) and secondly the remaining effects of lean and occlusion in the ortho-photos which will distort the classification results. These challenges were addressed by creating a 5 band image product (Near infrared, red, green, blue and height) directly from the Info Cloud through rasterization, thus generating a true ortho-photo, albeit with many voids at locations with tree lean. As the height information is derived directly from the imagery itself there are no registration issues and the classifier works very well with the newly created true ortho-photos increasing the overall success rate signficantly.

7. Conclusions and Outlook

A DSM derivation approach based on Semi-Global Matching (SGM) for ADS line scanner images was presented and compared DSM and DEM results with LiDAR. It was found that the SGM-derived surface strongly agrees with the LiDAR points. Based on high resolution ADS imagery, the increased point density reveals fine detail that may be difficult for LiDAR to capture. There are, however, significant differences inherent to the respective method – generally around trees and vegetation, where LiDAR in contrast to image-based SGM has the ability to penetrate to the ground, or measure the top more consistently. The apparent advantage of LiDAR to generate a bare earth model needs to be weighed against the much lower acquistion cost and the superior DSM product that embeds the full image radiometry in the point cloud in SGM.
In conclusion, we have shown that SGM-derived DEMs and DSMs are revealed to be an effective alternative to LiDAR, especially when high resolution is a requirement. While both data sets can generally be used for the purpose of orthoimage rectification, SGM is the preferred choice as it is based on the same data set – representing the same point in time, identical geometry and resolution – and saves the additional acquisition cost of LiDAR. SGM has been proven to be very successful in North West Geomatics' production and DSMs are generated virtually of all image projects by default. The high quality and the wealth of information in the final product has created much positive feedback, especially in industries that use both remote sensing and photogrammetry.
Based on this practical experience and upcoming needs, the presented SGM approach will continue to be refined and optimized for performance.

8. References

Alobeid, A., Jacobsen, K., Heipke, C., 2009. Building Height Estimation in Urban Areas from Very High Resolution Satellite Stereo Images. ISPRS Workshop: High-Resolution Earth Imaging for Geospatial Information, Hanover, Germany.

Gerke, M., 2008. Dense Image Matching in Airborne Video Sequences. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, China, Vol.

XXXVII, Part B3b, pp. 639-644.
Gehrke, S., K. Morin, M. Downey, N. Boehrer, and T. Fuchs, 2010: Semi-Global Matching: An Alternative to LiDAR for DSM Generation? Int. Arch. Phot. & Rem. Sens., Calgary, AB, 38(B1).

Heinrichs, M., Rohdehorst, V., Hellwich, O., 2007. Efficient Semi-Global Matching for Trinocular Stereo. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Munich, Germany, Vol. XXXVI, Part 3/W49A, pp. 185-190 (PIA07).

Hermann, S., Klette, R., Destefanis, E., 2009. Inclusion of a Second-Order Prior into Semi-Global Matching. 3rd Pacific-Rim Symposium on Image and Video Technology, Tokyo, Japan.

Hirschmüller, H., 2005. Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. Proc. IEEE Conference on CVPR, New York, New York.

Hirschmüller, H., 2008. Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, No. 2.

Hirschmüller, H., Mayer, H., Neukum, G., and the HRSC Co-Investigator Team, 2006. Stereo Processing of HRSC Mars Express Images by Semi-Global Matching. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Goa, India, Vol. XXXVI, Part 4.

Hirschmüller, H., Scharstein, D., 2007. Evaluation of Cost Functions for Stereo Matching. Proc. IEEE Conference on CVPR, Minneapolis, Minnesota.

Hirschmüller, H., Scholten, F., Hirzinger, G., 2005. Stereo Vision Based Reconstruction of Huge Urban Areas from an Airborne Pushbroom Camera (HRSC). Proc. 27th DAGM Symposium, Vienna, Austria, Vol. LNCS 3663, pp. 58-66.

Kato, A., Moskal, L. M., Schiess P., Calhoun, D., Swanson, M. E., 2010, True Orthophoto Creation Through Fusion of Lidar Derived Digital Surface Model and Aerial Photos, ISPRS TC VII Symposium – 100 Years ISPRS, Vienna, Austria, July 5–7, 2010, IAPRS, Vol. XXXVIII, Part 7A

Krauß, T., Lehner, M., Reinartz, P., 2008. Generation of Coarse Models of Urban Areas from High Resolution Satellite Images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Beijing, China, Vol. XXXVII, Part B1, pp. 1091-1098.

Leica Geosystems, 2009. Tech Note – Post Processing Times for ALS Data.

Morin, K., El-Sheimy, N., 2001a. A comparison of airborne laser scanning data adjustment methods. Proc. ISPRS WGII/2 Three-Dimensional Mapping from InSAR and LIDAR Workshop, Banff, Alberta, Canada.

Morin, K., El-Sheimy, N., 2001b. The Effects of Residual Errors in Airborne Laser Scanning Terrain Data on Ortho-Rectified Imagery. Proc. of Optical 3D Measurement Techniques V, Vienna, Austria.

Morin, K., El-Sheimy, N., 2002. Post-mission Adjustment Methods of Airborne Laser Scanning

Data. Proc. FIG/ASPRS Annual Conference, Washington, D.C.

Sun, H., Morin K., et al., 2006. IPAS – Leica Geosystems' High Accuracy GPS/IMU Integrated System for Airborne Digital Sensors. Proc. ASPRS Annual Conference, Reno, NV.

ALS monitoring of changes in forest biomass carbon storage

Chris Hopkinson[1,2], Laura Chasmer[2], Natascha Kljun[3], Eva van Gorsel[1], Harry McCaughey[4], Alan Barr[5], Andy Black[6]

[1]Centre for Marine and Atmospheric Research, CSIRO, Canberra, Australia
Hopkinsoncd@gmail.com
[2]Cold Regions Research Centre, Wilfrid Laurier University, Waterloo, ON. Canada
[3]School of the Environment & Society, University of Swansea, Swansea, UK
[4]Department of Geography, Queens University, Kingston, ON. Canada
[5]National Hydrological Research Centre, Environment Canada, Saskatoon, SK. Canada
[6]Faculty of Land & Food Systems, UBC, Vancouver, BC. Canada

Paper Number: #SL2012-038

Abstract

Multi-temporal ALS and field plot data are used to model changes in carbon storage over several years at jack pine and eucalyptus forest stands in Canada and Australia, respectively. Results are compared with continuous eddy covariance estimates of net ecosystem exchange and gross ecosystem production. Combined, these carbon monitoring techniques can yield improved partitioning of forest carbon storages and pathways through time.

1. Introduction

The United Nations Framework Convention on Climate Change (UNFCCC) is an international agreement requiring countries to report on forest carbon stocks and changes (IPCC 2006). Flux towers employing eddy covariance (EC) techniques (Baldocchi and Meyers, 1998) provide localised estimates of the carbon balance and require continuous operation to provide long term estimates. These data can be used to calibrate remote sensing data products (Heinsch et al., 2006) so that above ground carbon (AGC) changes can be reported over large scales. Airborne laser scanning (ALS) has been used to map above ground biomass (e.g. van Aardt et al., 2006; Hopkinson et al., 2011) and since data availability is increasing while costs are dropping, better understanding the capacity for ALS to monitor biomass and growth is of great interest (e.g. Hopkinson et al., 2008). Furthermore, CO_2 flux monitoring stations require continuous operation and frequent maintenance, creating uncertainty over long term data availability and a need for complementary monitoring methods.

We report on efforts to evaluate multi-year biomass monitoring at two sites: i) the Boreal Ecosystem Research and Monitoring Sites (BERMS) in Saskatchewan Canada, and ii) a Eucalyptus forest at Tumbarumba in New South Wales, Australia. We report on the ability to model changes in forest carbon storage at stands ranging in maturity (BERMS) and management (Tumbarumba), while comparing long term ALS biomass change estimates with flux tower records of gross ecosystem production (GEP) and net ecosystem exchange (-NEE) (-ve sign as atmosphere is sink).

2. Methods

2.1 Study areas

The BERMS area is located in the southern boreal forest of central Saskatchewan, Canada, experiences ~500 mm annual precipitation and a mean temperature of ~0.4 ^{o}C. The sites were established, in part, by the Boreal Ecosystems Atmosphere Study (BOREAS) in 1993 and continued from 2001 to 2011 as part of an effort to better understand the impacts of climate change and disturbance on the carbon cycle of terrestrial ecosystems (Fluxnet-Canada, Canadian Carbon Program) (Margolis et al., 2006). The BERMS sites include mature trembling aspen (*Populus*

tremuloides) and black spruce (*Picea mariana*), a jack pine (*Pinus banksiana*) chronosequence, peatland and forest fire sites. In this paper, four jack pine sites have been studied, due to the availability of multiple ALS datasets. The old jack pine (OJP) stand was cleared in 1914, whilst the other three stands were clear cut harvested in 1975 (HJP75), 1994 (HJP94) and 2002 (HJP02). These four stands are spread across an area of 5 km x 8 km, ranging in elevation from 500 m to 520 m with stem heights for the two oldest stands reaching ~ 20 m (Figure 1A, 1B).

The Tumbarumba study site is in the Bago State Forest of the southern tablelands of New South Wales, Australia. It has a moist temperate climate with annual precipitation ~ 1500 mm and a mean annual temperature of 8.0 ^{o}C (Leuning et al., 2005). The wet sclerophyll forest site ranges in elevation from 1175 m to 1325 m, is dominated by mature alpine ash (*Eucalyptus delegatensis*) with stem heights up to 50 m, drains a small headwater catchment and has undergone significant forest management in recent years (Figure 1C, 1D). The focus of the study is the area immediately surrounding a flux tower, which lies at the boundary of two stands that were operationally thinned in 1984 and 1985. Surrounding these stands (>500 m from the tower), commercial thinning operations have occurred prior to and following the installation of the EC flux tower in 2000.

2.2 Data collection

Summertime ALS data were captured over BERMS each August for 2005, 2008 and 2011 using a small-footprint multiple-discrete-return airborne laser terrain mapper (ALTM) 3100C (Optech Inc. Toronto) operated with comparable flight and sensor settings, resulting in an aerial sampling density >2 pts/m^2. Coincident with the 2005 ALS mission, species, height and diameter at breast height (DBH) were measured at 22 x 11.3 m (0.04 ha) radius permanent sample plots (PSPs) (8 for OJP and HJP75, 6 for HJP94). The plots were located in a radial pattern at 100 m and 500 m out from a flux tower located near the centre of each stand (Figure 1A). All stems down to 2 cm diameter were measured and all plot centres were surveyed with a dual frequency differential global position system (DGPS) to within 10 cm. Flux monitoring towers were operational at the sites from ~2000 to 2008 (more recent data were collected but are unavailable at present).

At Tumbarumba, summertime ALS data were captured in December 2001 (Optech ALTM 2050), November 2009 (Riegl LMS-Q560) and January 2011 (Leica ALS60) (January 2011 is assumed to represent conditions at end of 2010). Unlike BERMS, the ALS data were not collected using equivalent acquisition configurations. Based on scan geometry, discrete return properties and sampling density (>2pts/m^2), the Optech ALTM and Leica ALS datasets collected nine years apart at ends of 2001 and 2010 were the most comparable. Within approximately 1 km of the flux tower, there are 30 x 17.8m radius (0.1 ha) PSP). In December 2009 (~ one year prior to the 2010 ALS survey), 20 of the PSP centres were surveyed with a single frequency DGPS to within 2 m, and DBH values for all stems were recorded. NEE and GEP data are available for the entire period.

2.3 Analysis

Field data were used to estimate stem- and plot-level above ground biomass (AGB) using established species-specific allometric equations. For the jack pine stands, stem height and DBH data were used (Lambert et al., 2005), while for the eucalypt stands DBH data alone were applied to a native sclerophyll forest equation (Keith et al., 2000). No plot data were collected at HJP02 for this study but biometric data were collected in 2004 by Theede (2007) and are available on the Fluxnet-Canada Data Information System (DIS) (http://fluxnet.ccrp.ec.gc.ca/e_DataAccess.htm). Consequently, AGB at HJP02 was estimated from data collected in 2004. Plot-level biomass data were used to test ALS models of biomass by extracting coincident point cloud data and running regression models against many commonly used point cloud metrics. All ALS points within a plot were normalised by computing elevation residuals relative to a ground-classified digital elevation model (DEM) so that heights were reported relative to ground. The following metrics were tested for all returns and all returns above 0.2 m (to remove ground influence): mean, maximum, standard deviation, inter-quartile range (IQR), percentiles [P25, P50, P75, P90, P95, P99] and the following ratios: i)all returns / all

returns above 1.5m, ii) all returns / all returns above average height and iii) all returns / all first returns above 1.5m. Single variable linear, power, log and exponential models were tested.

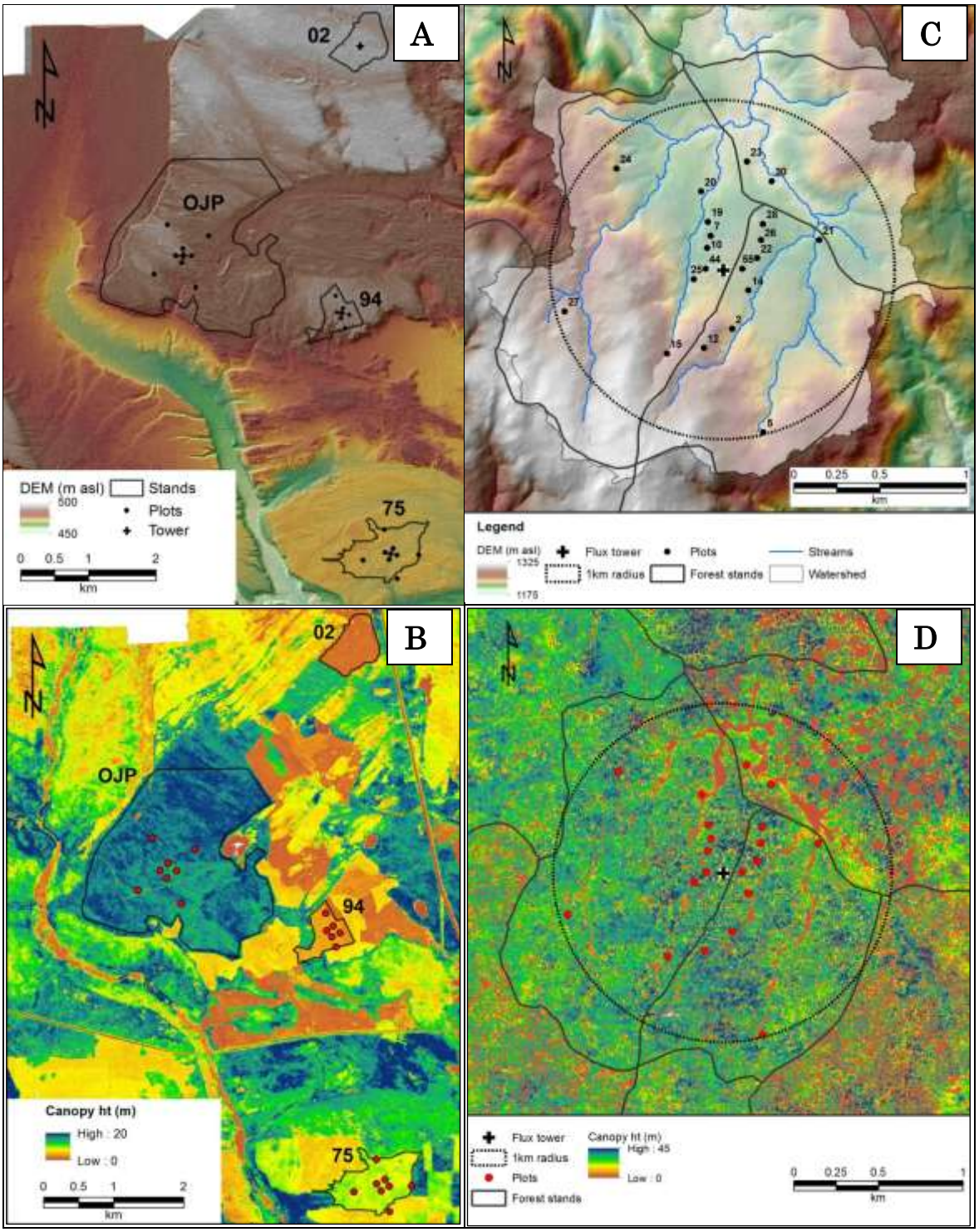


Figure 1: A) DEM, forest stands and plot locations for BERMS; B) Canopy height model (CHM) for BERMS illustrating jack pine stand heights in 2011; C) Tumbarumba DEM, forest stands, and plots with 1 km radius around tower to illustrate proximity to stands and headwater catchment; D) Tumbarumba CHM in 2010 illustrating variations in canopy openness associated with different stand management practices.

An optimal ALS model was chosen, based on predictive capability and consistency of its application

to different datasets; i.e. a model could theoretically reproduce plot-level biomass estimates for the year of coincident plot and ALS sampling yet when applied to other years produce unrealistic or incomparable biomass estimates across the same landscape. Aside from differences in canopy sampling due to variations in laser pulse strength or geometry (Hopkinson, 2007), differences in understory, ground cover, leaf area index or surface moisture could potentially alter the point cloud frequency distribution shape and therefore the efficacy of the chosen model. Temporal plot data were not available for this test of comparability, so judgement was exercised based on years of field observations and knowledge of canopy conditions and management treatments at each site. For these reasons, a simple univariate model was preferred over a more sophisticated model adopting multiple variables or curve fitting techniques that could easily be trained to a dataset that itself might contain significant uncertainty. The assumption is that a simpler technique might not produce the highest accuracy model but it should be more robust for comparative analyses across a range of canopy heights and densities (e.g. Hopkinson et al., 2004).

Once an ALS AGB model was calibrated for BERMS and Tumbarumba, it was applied to the ALS point clouds surrounding each site. ALS model data were gridded to 20 m x 20 m cell arrays due to the BERMS training data that were extracted from 400 m^2 areas and a decision by AusCover to develop national ALS variable datasets at this raster resolution. This resolution was also convenient in that it ensured there were sufficient ALS points in each cell to generate a complete frequency distribution, while mitigating slight positional uncertainties in any of the datasets compared. AGB was converted to AGC_{ALS} (above ground carbon estimated from ALS) by assuming 50% of dry biomass is C and thus applying a multiplier of 0.5 (Atjay et al. 1977) to the model grids.

An expansion factor *e* (i.e. a multiplier greater than 1) was applied to AGC_{ALS} grids to derive $eAGC_{ALS}$ and add in a below ground biomass (BGB) component, so that tower-based productivity estimates would be more directly comparable with the biometric estimates derived from ALS. Dead biomass and other carbon pools were not included in this biomass expansion, as the intent was to map long-term productivity (growth). At BERMS, BGB *e* factors were derived from the biometric database on the Fluxnet DIS and ranged from 30% at the mature OJP site to 100% (or a doubling of AGC) at the HJP02 site. No data were available to derive an accurate expansion factor for the mature eucalypt stands, so a generic *e* factor of 20% was applied (Greenhouse Challenge, 1998).

At BERMS and Tumbarumba, stand- and plot-level estimates of $eAGC_{ALS}$ were estimated for each year of ALS data. By summarising the plot-level estimates, it was possible to infer whether or not the plots accurately captured the spatial variability of biomass within the stands. The total change in above ground carbon ($\Delta eAGC_{ALS}$) was calculated by differencing the grid arrays and summarising by stand, plot and, in the case of Tumbarumba, the 1 km radius area surrounding the flux tower. These results were then compared with available GEP records (BERMS), and -NEE and GEP (Tumbarumba) to assess the relative magnitudes of biometric ALS estimates of forest carbon storage change with those obtained from atmospheric flux measurements.

3. Results and discussion

3.1 AGC_{ALS} model

Plot-level AGB derived from allometric equations ranged from a minimum of 0.8 t C/ha at HJP94 to a maximum of 50.9 t C/ha at OJP, and 30.8 t C/ha to 305 t C/ha over Tumbarumba stands. At BERMS, the field mensuration and allometric equations contained more information and resulted in reduced uncertainty in the biomass training data than at Tumbarumba. Consequently, iterative model testing was performed on the BERMS 2005 data first and when suitable metrics identified, testing was then carried out on the Tumbarumba 2010 data. Plot-level model results for BERMS are presented in Table 1. The best fit regression model using a single independent variable at BERMS was a power function of the average height (Avg_{all}) of the 'all return' point cloud distribution ($r^2 = 0.98$, RMSE = 2.4 t C/ha, n = 22). However, this model was considered unsuitable for four reasons: i) it did not pass through the origin and it is to be expected that at zero AGB, there should be no point cloud data (apart from noise) overlying the ground surface; ii) outside the height range of the training dataset, biomass estimates

increased unrealistically for small increases in Avg_{all}; iii) when applied to 2008 and 2011 datasets, the grid comparisons demonstrated unrealistic gains and losses in biomass at locations where it is known biomass was accumulating at a relatively steady state; iv) Avg_{all} was sensitive to outliers associated with scan artefacts in swath overlap regions. From Table 1, it is also clear that simple linear models with a more physically realistic zero origin produced model results that were not appreciably inferior to non-linear models. Avg_{all} still produced the best fit model to the training data (i.e., $r^2 = 0.97$) but still suffered from artefacts and unrealistic comparison results when applied to subsequent datasets.

Table 1: Plot-level AGC_{ALS} model test results at the BERMS jack pine stands.

Point cloud metric	All returns		All returns > 0.2m	
	linear model through origin (r^2)	non-linear model (r^2)	linear model through origin (r^2)	non-linear model (r^2)
P25	0.41	0.53	0.45	0.79
P50	0.09	0.43	0.92	0.96
P75	0.96	0.97	0.94	0.94
P90	0.95	0.97	0.93	0.93
P95	0.94	0.96	0.92	0.93
P99	0.92	0.92	0.91	0.91
Max	0.9	0.91	0.9	0.91
Avg_{all}	0.97	0.98	0.9	0.93
SD	0.95	0.96	0.91	0.91
IQR_{all}	0.96	0.97	0.56	0.87
All/1.5	0.62	0.91	0.6	0.88
All/avg	0.34	0.75	0.24	0.81

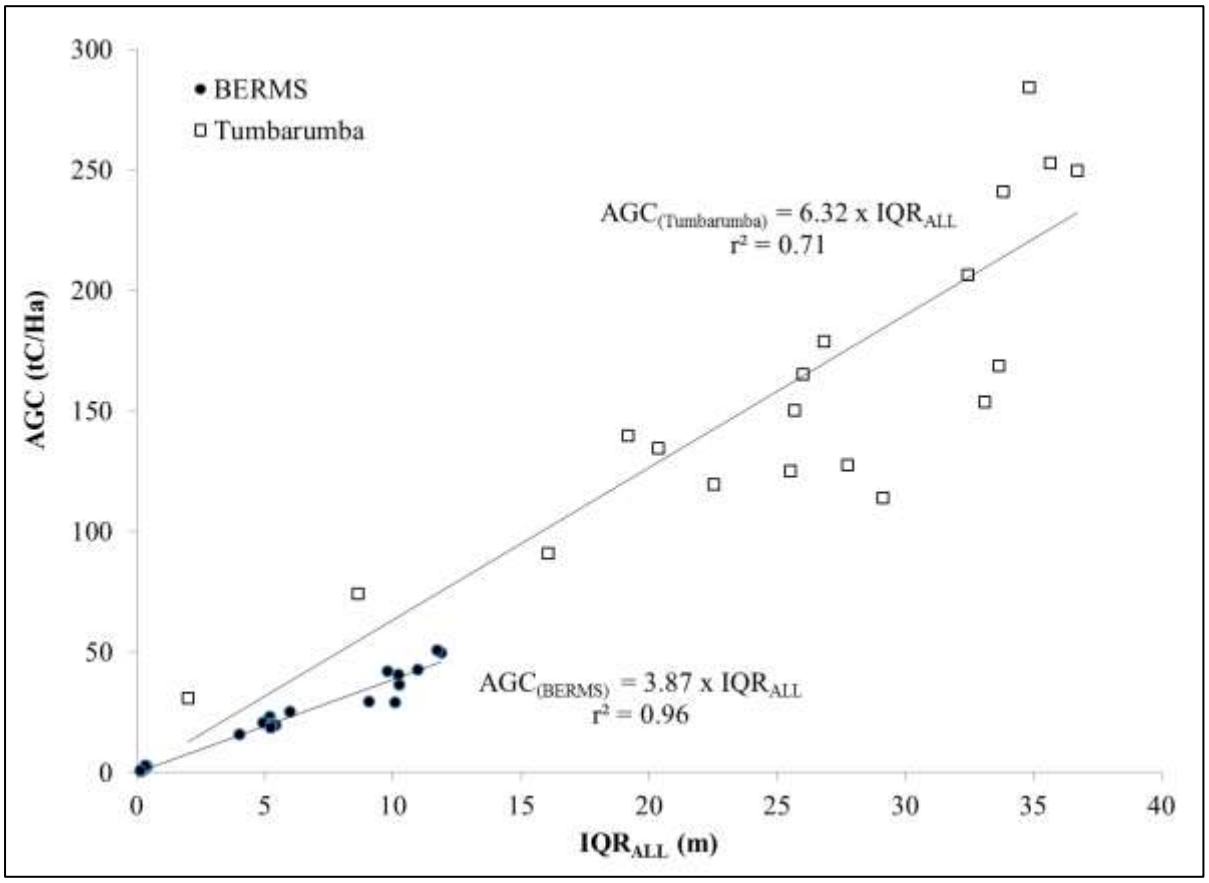


Figure 2: Linear best fit models of AGC_{ALS} from IQR_{all} at BERMS and Tumbarumba.

Interquartile range (IQR_{all}) was found to produce the best compromise between accurate model results (r^2 0.96, RMSE = 3.2 t C/ha, n = 22), and realistic temporal comparisons. IQR_{all} was thought to

provide a superior variable for biomass prediction because it represents a surrogate measure of overall point cloud height and density without being over-sensitive to the tails of the distribution. IQR_{all} was also found to provide the best fit model at Tumbarumba (r^2 = 0.71, RMSE = 37.6 t C/ha, n = 20) (Figure 2). That r^2 was lower than at BERMS is likely due to the plot biomass estimates being generated from DBH only and the adoption of a more generic allometric equation. Nonetheless, it is encouraging that at two sites with very different canopy structure and ALS point cloud attributes, biomass could be reasonably accurately predicted using a simple linear function of the point cloud IQR. The multiplier was lower at the jack pine sites and this is likely because of the shorter trees, higher stem densities and more clumped canopy biomass characteristic of the boreal forest.

3.2 Stand-level Δ $eAGC_{ALS}$

Stand- and plot-level estimates of $eAGC_{ALS}$ and $\Delta eAGC_{ALS}$ (Figure 3A) at BERMS were similar and within the range of observed standard deviations (Table 2). However, stand-level biomass estimates for OJP were almost 10% greater than those calculated from the modeled results at the plot locations alone. This discrepancy is because the plots were located around the flux tower, which is in an area of relatively short canopy (Figure 1B) (and therefore low biomass) when compared to the rest of the stand. Consequently, while the plots provide a reasonable estimate of biomass in the area around the tower, the estimate aggregated from the raster cells within the stand is probably a more reliable indicator of more widespread jack pine conditions for a stand of this age and in this region.

Table 2: BERMS stand- and plot-level $eAGC_{ALS}$ estimates and the associated change between years surveyed. Also shown is the ratio of $\Delta eAGC_{ALS}$ to mean GEP for the study period (GEP periods vary due to availability)

Stand-level	Mean $eAGC_{ALS}$ (t C/ha) (SD)			$\Delta eAGC_{ALS}$ (t C/ha)			$\Delta eAGC_{ALS}$ / GEP 2005-2011
	2005	2008	2011	2005 - 08	2008 - 11	2005 - 11	%GEP (estimated.)
HJP02	0.29 (0.20)	0.37 (0.51)	1.1 (0.78)	0.1	0.7	0.8	1%
HJP94	1.9 (0.70)	7.5 (2.8)	14.7 (3.2)	5.6	7.3	12.8	36%
HJP75	31.1 (2.1)	33.2 (2.1)	37.1 (2.3)	2.1	3.9	6.0	11%
OJP	59.1 (7.6)	59.7 (6.8)	60.7 (7.4)	0.5	1.0	1.5	4%
Plot-level							
HJP94	1.8 (0.70)	5.9 (3.9)	12.6 (5.2)	4.2	6.7	10.8	30%
HJP75	30.2 (3.0)	32.4 (2.7)	36.2 (2.7)	2.2	3.9	6.1	11%
OJP	53.7 (4.6)	54.7 (4.4)	55.4 (4.9)	1.0	0.7	1.7	5%

NEE and GEP were not available for all of 2005 to 2011 at all the BERMS jack pine sites. Mean annual $\Delta eAGC_{ALS}$ was therefore compared to mean annual GEP for the years of data availability at each site to derive an estimate of 'GEP efficiency'; i.e. an estimate of the proportion of GEP (photosynthesis) that has been utilized within each stand to increase its total living biomass (Table 2). From this we infer that GEP efficiency at the jack pine stands is smallest at the youngest and oldest sites (HJP02 and OJP, respectively) and is highest at the immature HJP94. These estimates suggest annual woody biomass increases are small relative to seasonal biomass production and biomass replacement at young sites where there is high mortality and in mature sites where a high proportion of GEP goes into maintaining already high levels of biomass.

Table 3: Tumbarumba $eAGC_{ALS}$ estimates and the associated change.

		Mean $eAGC_{ALS}$ (t C/ha)				$\Delta eAGC_{ALS}$ (t C/ha)
		2001		2010		2001-10
Location	Area (ha)	Avg	SD	Avg	SD	
Tower stands	341	184	79	207	61	23
1 km radius around tower	314	179	86	193	77	14
Field plots	2	168	53	176	35	8

At Tumbarumba, stand, 1 km tower radius and aggregated field plot $eAGC_{ALS}$ estimates varied by 16

t C/ha or < 10% (2001) and 31 t C/ha or <18% (2010) and these differences were within the range of observed standard deviations (Table 3). In both years, plot estimates were the lowest, with stand estimates being the highest and the 1 km tower radius area being intermediate. The implication of these relatively small differences in biomass estimate is clear in the $\Delta eAGC_{ALS}$ estimates, where the plot-based estimate is 57% that of the tower radius and 35% that of the stands surrounding the tower. These observations illustrate large spatial variability in biomass change across this eucalyptus forest ecosystem (Figure 3B) and an inherent challenge in identifying the ideal spatial domain for ecosystem characterization. This is particularly important when relating biometric estimates of biomass change to tower-based estimates of ecosystem productivity and thus highlights the need to integrate spatial ALS and point-based EC data using flux footprint models (Kljun et al., 2004; Chasmer et al., 2008).

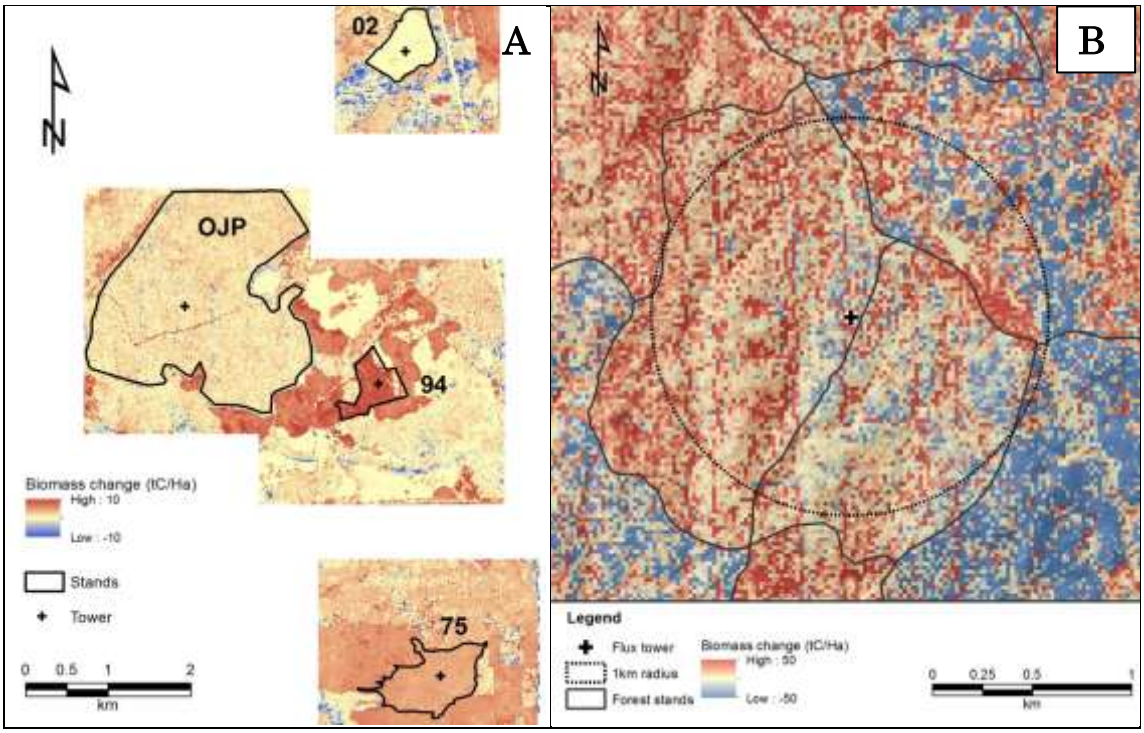


Figure 3: $\Delta eAGC_{ALS}$ at BERMS (6 yrs) and Tumbarumba (9 yrs). Biomass losses at both sites are primarily due to commercial thinning operations, while gains are due to natural stem growth.

In the undisturbed stand area where the EC flux tower is situated, $\Delta eAGC_{ALS}$ indicates there has been ~23 t C/ha net biomass accumulation over nine years. This represents ~46% of cumulative -NEE (50 t C/ha) and ~10% of the cumulative annual GEP (232 t C/ha). Similar to the older BERMS OJP and HJP75 stands, this suggests that ~90% of the CO_2 uptake is used to support seasonal growth that subsequently dies and in replacing biomass lost from mature trees. Over the longer term, the -NEE (net CO_2 uptake) does not balance the ALS-modeled change in biomass. Some of this imbalance is likely due to spatial uncertainty in the tower footprint, e factor assumption-errors, plot-level allometric biomass estimates or due to inconsistency in the two ALS datasets. (Some imbalance is expected due to soil organic matter decomposition and heterotrophic respiration but this loss would act to close the gap we see not widen it). More effort is required to understand and mitigate these measurement uncertainties. If indeed, however, there is a large difference between ALS-modeled biomass accumulation and tower-based estimates of NEE, this suggests either a large and growing storage of CO_2 in the ecosystem or export out of the ecosystem via a pathway that is unaccounted for.

4. Conclusions

A straightforward ALS-based approach to monitoring the carbon in living tree biomass in managed

forests has been presented. By comparing these data with EC flux-based estimates of -NEE and GEP, it is possible to partition ecosystem production into growth components associated with: i) sustained biomass accumulation (of interest from a commercial growth and yield perspective); and ii) seasonal biomass cycling and replacement. More research is needed to understand long term CO_2 storages and pathways within the boreal jack pine and temperate eucalyptus ecosystems studied, and the uncertainties propagated during the ALS biomass modelling stages adopted in this study.

Acknowledgements

David Colville, Jorg Hacker, Arantxa Cebello, Darius Culvenor, Chris Beasy, Allyson Fox, Rebecca Trevithick, AusCover, Fluxnet-Canada, OzFlux, Canada Foundation of Innovation. Natural Environment Research Council UK, NE/G000360/1, National Centre for Earth Observation mission support 2009

References

Atjay, G.L., Ketner, P., Duvigneaud. P. 1977, Terrestrial primary production and phytomass. p. 129-181. In B.E. Bolin, T. Degen, S. Kempe and P. Ketner (ed.) The global carbon cycle. John Wiley, New York

Baldocchi, D., and Meyers, T. 1998, On using eco-physiological, micrometeorological and biogeochemical theory to evaluate carbon dioxide, water vapor and trace gas fluxes over vegetation: a perspective. Agric. For. Meteorol. 90, 1-25

Chasmer, L., Kljun, N., Barr, A., Black, A., Hopkinson, C., McCaughey, H. and Treitz, P. 2008, Vegetation structural and elevation influences on CO_2 uptake within a mature jack pine forest in Saskatchewan, Canada. *Canadian Journal of Forest Research*. 38, 2746-2761.

Greenhouse Challenge. *Greenhouse Challenge Vegetation sinks Workbook*. Canberra: Australian Greenhouse Office; 1998.

Heinsch, F.A., Zhao, M., Running S.W. et al. 2006, Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. *IEEE Trans. Geosci. Rem. Sens.* 44, 1908-1925.

Hopkinson, C. 2007, The influence of flying altitude and beam divergence on canopy penetration and laser pulse return distribution characteristics. *Canadian Jnl of Remote Sensing*, 33, 312-324.

Hopkinson, C. Chasmer, L., Hall, R.J., 2008. The uncertainty in conifer plantation growth prediction from multitemporal lidar datasets. *Remote Sensing of Environment*. 112 (3)

Hopkinson, C., Lim, K., Chasmer, L., Treitz, P., Creed, I. F., Gynan. C., 2004, Wetland grass to plantation forest - estimating vegetation height from the standard deviation of lidar frequency distributions. *Proceedings of the ISPRS working group VIII/2,'Laser-Scanners for Forest and Landscape Assessment* Freiburg, Germany 03-06 October 2004 ISPRS 36, 8/W2

Hopkinson, C., Colville, D., Bourdeau, D., Monette, S., Maher, R. 2011., Scaling plot to stand-level lidar to public GIS data in a hierarchical approach to map the biomass of Nova Scotia. *Proceedings of the SilviLaser 2011 Conference*, Oct. 16-20, Hobart, Tasmania.

IPCC, 2006, 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds). Published: IGES, Japan.

Kljun, N., Calanca, P., Rotach, M., and Schmid, H. 2004, A simple parameterisation for flux footprint predictions. Boundary-Layer Meteorology, 112, 503–523.

Keith, H., Barret, D., Keenan, R., 2000, *review of allometric relationships for estimating woody biomass for New South Wales, the Australian Capital Territory, Victoria, Tasmania & South Australia; National carbon accounting system tech. report No. 5B*. Australian Greenhouse Office, Canberra, Australia. 121pp.

Leuning, R., Cleugh, H.A., Zegelin, S., Hughes, D., 2005, Carbon and water fluxes over a temperate Eucalyptus forest and a tropical wet/dry savannah in Australia: measurements and comparison with MODIS remote sensing estimates. *Agriculture and Forest Meteorology,* 129, 151-173.

Margolis, H., Flanagan, L., Amiro, B. 2006, The Fluxnet-Canada Research Network: Influence of climate and disturbance on carbon cycling in forests and peatlands, *Agric. and Forest Meteorol*. 140, 1-5.

Theede, A. 2007. *Biometric and eddy covariance estimates of ecosystem carbon storage at two boreal forest stands in Saskatchewan: 1994-2004*. Unpublished Masters of Science thesis. University of Saskatchewan. 129 pp.

van Aardt, J., Wynne, R., Oderwald, R., 2006, Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per segment basis. *Forest Science*, 52, 636-649.

Robust characterization of forest canopy structure types using full-waveform airborne laser scanning

R. Leiterer[1], F. Morsdorf[1], M.E. Schaepman[1], W. Mücke[2], M. Hollaus[2], N. Pfeifer[2]

[1]Remote Sensing Laboratories, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland, Email: rleitere, felix.morsdorf, michael.schaepman@geo.uzh.ch
[2]Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Gusshausstrasse 27-29, 1040 Vienna, Austria, Email: wm, np, mh@ipf.tuwien.ac.at

Paper Number: SL2012-041

1. Introduction

Forests cover almost one third of the total land surface of the Earth. Therefore, they play a pivotal role in the global biogeochemical and -physical cycles between atmosphere and the land surface (Ross 2011; Betts *et al*. 2001). Understanding, assessing and quantifying forest ecosystem goods and services and their underlying processes (De Groot *et al*. 2002), helps to project the development of biogeochemical cycles under changing climate conditions (Jonsson and Wardle 2010; Sierra *et al*. 2009) and to develop sustainable management strategies (Purves and Pacala 2008). Particularly the complex three-dimensional distribution of geometric objects and their topology within forests canopies, here termed *forest canopy structure* (Nadkarni *et al*. 2008; Disney *et al*. 2006), influences the fluxes of energy and matter between the atmosphere and forests (Xue *et al*. 2011; Yang and Friedl 2003) and is one of the critical variables to determine forest stand resistance to disturbances and to estimate the conservation potential for biodiversity (Kayes and Tinker 2012; Lindenmayer *et al*. 2006).
Assessing forest canopy structure is difficult: conventional fieldwork is time-consuming, subjective and mostly limited in its spatial extent (Foody 2010; Haara and Leskinen 2009; Strand *et al*. 2002), whereas traditional remote sensing methods are lacking information in the vertical dimension (Jones *et al*. 2012; Hall *et al*. 2011; Roberts *et al*. 2007). Airborne laser scanning (ALS) systems have been shown to be suitable for providing not only horizontal information on the forest canopy structure, but also explicit vertical information due to the canopy penetration of the emitted signal (Kaartinen *et al*. 2012, Leeuwen and Nieuwenhuis 2010). Canopy structure metrics derived by ALS include geometric variables such as canopy height, canopy volume and canopy base height, as well as biophysical variables such as the Plant Area Index (PAI) or the canopy cover (Hilker *et al*. 2010; Morsdorf *et al*. 2009). However, existing approaches mostly include manual processing steps or need additional data about stand characteristics (e.g. tree species, age) (Antonarakis *et al*. 2011; Korpela *et al*. 2010; Kim *et al*. 2009). Therefore, a robust and transferable method is basically needed to provide a more efficient monitoring of forest canopy structure, as well as to improve the robustness and reliability of derived structure variables as input for environmental modeling, e.g. for dynamic global vegetation models or forest gap models.
To meet these requirements, we utilized the concept of forest canopy structure types (CSTs). Each specific CST represents a unique set of forest canopy structural variables. Thus, areas with the same CST are assumed to be homogenous in terms of forest canopy structure. We developed a robust and physical based method using full-waveform ALS data under leaf-on/leaf-off conditions in a deciduous dominated forest stand to extract a set of forest structure variables (crown and canopy dimensions, tree position, tree types, and occurrence of understory) as input for the CST derivation. For the validation, we used Digital Hemispherical Photography (DHP), Terrestrial Laser Scanning (TLS), and forest inventory data.

2. Method

Forest structure variables characterize canopies at different spatial scales: for example on individual tree level (e.g. crown length), on stand level (e.g. mean canopy height) or on landscape level (e.g. proportion of canopy cover) (Shugart *et al*. 2009). In this study, forest canopy structure will be investigated on the individual tree level and based on a regular grid, representing the stand level.

2.1 Study area and data

In the approach presented in this study, we use full-waveform ALS data in a 300 x 300 m test site. The data acquisition was performed under leaf-on (using RIEGL's© LMS-Q680i scanner) and leaf-off (using RIEGL's© LMS-Q560 scanner) canopy conditions in a mainly semi-natural, deciduous-dominated forest stand in Laegeren (Swiss Jura; 47°28'N, 8°21'E), yielding in two independent datasets. The used sensor specifications are summarized in Table 1 and described more in detail in Wagner *et al*. 2008 and RIEGLs (2012) specific sensor documentations.

Table 1: Used sensor specifications of RIEGL's laser scanner LMS-Q560 and LMS-Q680i.

	LMS-Q560	LMS-Q680i
pulse repetition rate [Hz]	200 000 Hz	
scan angle [deg]	± 15 deg	
mean operating altitude above ground [m]	500 m	
date of acquisition	10.04.2010	01.08.2010

The dense forest stand is characterized by steep topographic relief (slope up to 60°) and high species diversity (mostly *Fagus sylvatica (L.)*, *Picea abies (L.) Karst*, *Fraxinus excelsior (L.)* and *Acer pseudoplatanus (L.)*), age (55-160 years), and diameter distribution (7-120 cm) (Eugster *et al*. 2007). For the investigated area, an extensive set of ground based reference data is available, mainly measured during field campaigns in September 2011: DHPs, TLS, and forest inventory data. All field measurements were geo-referenced and co-registered based on traditional terrestrial land surveying by a total station and a GNSS real-time kinematic system.

2.2 Data pre-processing

The benefit of full-waveform data is the approximation of the entire backscattered signal by digitization, which facilitates the extraction of additional features of each reflecting object within the ALS footprint. To detect and extract specific object reflections, Gaussian pulse estimation was applied in order to obtain representative echo descriptions. In particular, this includes the derivation of the point cloud with its basic and established geometrical characteristics of reflectors but additionally the physical description of each reflector with information such as amplitude, width and intensity of each specific echo.
In a first step, we extracted the ground returns from the point cloud using the single and last echoes, their geometrical characteristics and echo width information (Mücke *et al*. 2010). Based on the method by Evans and Hudak (2007) we developed a new adaptive multi-scale filter algorithm. As part of the filtering, a kernel based query was applied to the selected ground return echoes to detect areas with high deviations in height values (≥ 100 % slope). Within these areas a combination of optimized spline function analyses (amount of local maxima) and a scale dependent multi-point triangulation (in a 3x3 and/or 5x5 kernel) was applied to exclude non-ground points from the point cloud. The spline function approach allows a reliable distinction between height deviations caused by the steep terrain and the less continuous height deviations caused by artificial objects or dense vegetation. The remaining points were interpolated applying an ordinary kriging to a 1 x 1 m digital terrain model (DTM). Additional,

a 1 x 1 m digital surface model (DSM) was processed using the first echo reflections and their corresponding echo width. Afterwards, the canopy height model (CHM) was calculated from the difference between DSM and DTM. Finally, for each point of the point cloud we determined the height above ground as well as the according DTM, DSM and CHM values.

2.3 Extraction of structure variables on tree level

The characterization of the crown structure of individual trees requires the segmentation of the full point cloud into specific point clouds of the single trees. As input for the point cloud segmentation a set of seed points is necessary, representing the position of the individual trees. To detect understory trees as well, we applied an iterative, three-dimensional grayscale dilation on the point cloud based on an ellipsoid-shaped structuring element S with a pre-defined domain D_S (Adams 1993). Accordingly, all resulting local maxima were used as initial seed points for a standard k-means clustering. The clustering with an Euclidean metric favors ball shaped clusters in a three-dimensional feature space. Due to the high variations in the vertical extent of the point cloud, a scaling of the height values (z-coordinates) was applied using the ratio of the respective CHM value to the height above ground of each specific point. Based on the resulting individual point clusters, we calculated the alpha shape (Vauhkonen *et al*. 2009) to be able to derive additional crown specific variables such as crown volume or crown diameter.
For all points within the individual alpha shapes, we investigated the variations in the point distribution between the leaf-on and leaf-off acquisitions (e.g. percentage distribution in vertical extent). Based on the differences (using a significance level of 5 %), we distinguished deciduous from coniferous trees. This attribute as well as the information about the unique point cloud membership was added to each point within the crown.

2.4 Determination of canopy structure types (CSTs)

To derive CSTs, we analyzed the vertical stratification and properties of the specific point clusters based on Cartesian grids with 1 x 1 m, 5 x 5 m and 50 x 50 m pixel size, respectively. Depending on the respective spatial scale, we extracted the common canopy variables mean canopy height, canopy cover, mean length of live canopy, mean height to canopy base, the occurrence and height of understory and the foliage distribution as well as their variations within the pixel (except for the 1 x 1 m grid). This includes information about the occurrence of various canopies in a vertical column and the species composition in terms of coniferous and deciduous trees. Figure 1 shows the concept of the vertical forest canopy stratification for the 5 x 5 m grid.

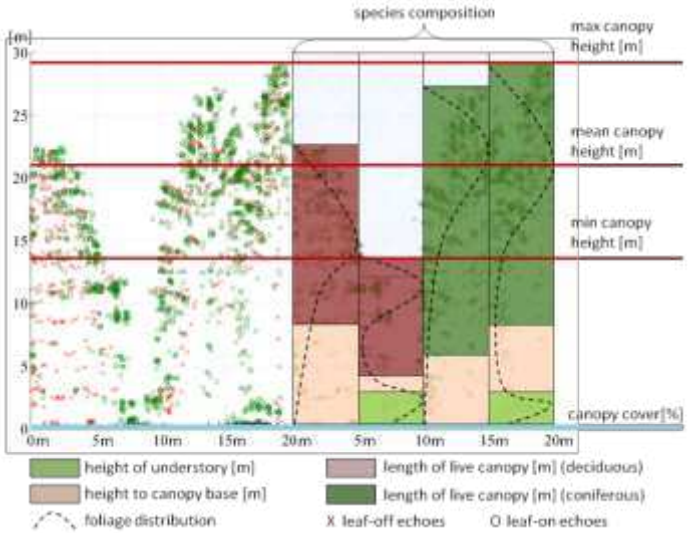


Figure 1: Example for the grid based vertical stratification of forest canopy structure (cell size 5 x 5 m).

For the 50 x 50 m grid, we applied the classic index by Clark and Evans (1954) for the statistical determination of distribution patterns based on the individual trees extraction (random, uniform, and clumped) as well as determined stand level parameters such as stand density and stand height. The resulting grid with the multi-dimensional feature space was then classified into a pre-defined amount of unique CSTs, representing structural homogeneous forest areas.
To evaluate the applicability of the derived CSTs for the estimation of bio-physical variables, we enhanced the method of Morsdorf *et al*. (2006) to calculate the PAI, the fraction of absorbed photosynthetically active radiation (fAPAR) and the canopy cover, utilizing the amount of vertical layers, the foliage distribution and the tree type of the specific CSTs. For the validation of the estimated parameters, both point measurements with DHPs and plot-wise TLS measurements were carried out, resulting in grid based PAI, canopy cover and fAPAR layers.

3. Results and discussion

The reconstruction approach on tree level results in a complete 3D representation of the forest area and the CST classification. Figure 3 shows a visualization of the 3D-scene and the result of the CST classification for the 5x5 m grid.

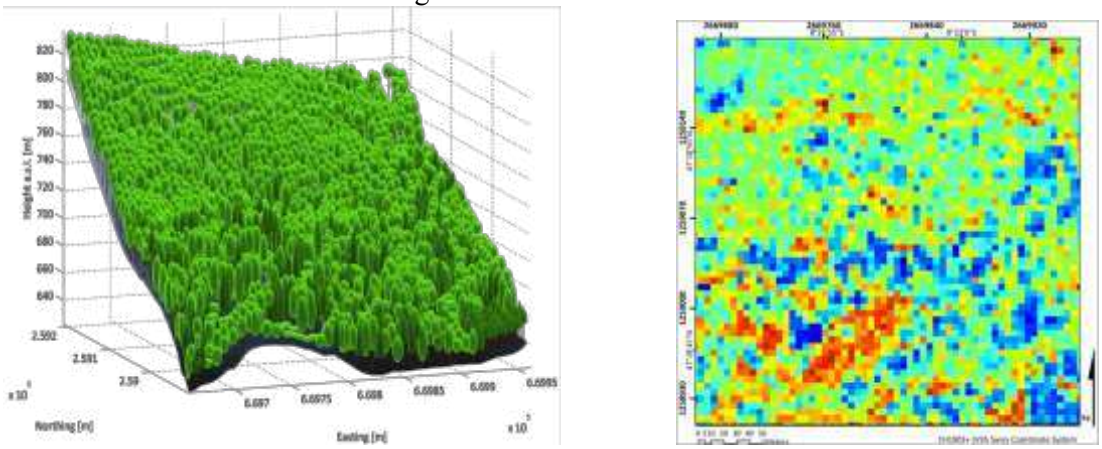


Figure 3: Visualization of the reconstructed three-dimensional forest scene and the 5x5 m CST map.

In the 3D representation, the complex alpha shapes of the canopies are simplified using geometrical primitives depending on the tree species (ellipsoids for deciduous; paraboloids for coniferous); whereas the stems are represented by uniform cylinders and placed at the specific local maximum position. In the map of the CSTs, each color is representative for a unique combination of the derived canopy structure variables.
The validation of the tree/ stem position was based on TLS measurements for two 30 x 30 m plots and leaf-off/on ortho-images outside these areas. The commission and omission errors for the delineation is 5.2% and 13.1%, respectively, whereby the omission errors have mainly been caused by the large amount of clustered and multi-stemmed trees due to former coppicing activities (Van Calster *et al*. 2008). The connected canopies of these clusters are even with field measurements nearly inseparable and show a similar shape as the crown of an individual tree. As the applied tree delineation method is based on the assumption, that each crown represents an individual tree/ stem, we underestimate the amount of trees/ stems particularly in the old beech stands. Thus, the estimation of the related biomass and stem volume information as well as the determination of the stand density are likely to be less accurate and were not further investigated in the context of this study. The distribution patterns of the individual trees were compared with the reference data based on a grid representation of the three classes “random”, “uniform”, and “clumped”, resulting in a mean r^2 of 0.56. However, this result is subject to the same limitations as the detection of the individual trees.
The cross-comparison of the crown dimensions with the TLS measurements, ortho-images and forest inventory data shows a high consistency in the horizontal dimension, whereas a

quantitative validation of the horizontal crown dimensions and the crown volume/ crown surface is difficult and involves a high level of uncertainty (cf. tree detection). The results of the extraction of the vertical extent is shown in Table 2 for the individual tree level and the canopy level based on the specific CSTs, with slightly better results for the CST based dimension variables. The occurrence of understory was detected with an accuracy of 78% with a mean error in the vertical extent of 1.2 m.

Table 2: Accuracy assessment for the extraction of vertical crown/ canopy dimension variables.

mean error of	**individual tree level**	**canopy level**
crown height / mean canopy height [m]	0.6	0.5
length of live crown / mean length of live canopy [m]	2.8	1.2
height to crown base / mean height to canopy base [m]	2.6	1.6

The classification in coniferous and deciduous trees was compared to the forest inventory data and the ortho-images utilizing the common statistics metrics overall accuracy (OA), users accuracy (UA), producers accuracy (PA) and the kappa coefficient (Liu *et al*. 2007) (Table 3). In comparison to existing studies for species distinction (e.g. by Kim *et al*. 2009), we achieved similar accuracies using only the leaf-on/leaf-off point distribution varieties. Accordingly, the species distribution derived from the CSTs differs 10 to 15 % to the reference data. The major sources of uncertainty in the test area are the misclassification of *Larix decidua (Mill.)* and areas with structure changes in the canopy caused by wind and snow induced damages.

Table 3: Confusion matrix of the distinction between coniferous and deciduous trees.

	coniferous (ALS derived)	**deciduous** (**ALS derived**)	**PA [%]**
coniferous (reference)	430	79	84.5
deciduous (reference)	120	1300	91.5
UA [%]	78.2	94.3	
OA [%] / Kappa coefficient			89.7 / 0.74

The CST derived bio-physical variables show high correlations to the reference data derived from the DHPs (PAI r^2 = 0.57, fAPAR r^2 = 0.64, canopy cover r^2 = 0.78). However, the statistical base for the bio-physical variables of only 18 reference measurements is insufficient for a representative and quantitative validation. Therefore, we need to extend the ground measurements to a larger area within the future work. Nevertheless, the results are promising considering the robust, physical based approach of the CSTs. The derived foliage distribution information could not be validated yet directly. For this purpose a dense vertical sampling of the full canopy (e.g. leaf area density) is necessary, which will be carried out in future studies.

4. Conclusion

In this study, we developed a robust and transferable method for a physically-based extraction of canopy structure variables on the individual tree level as well as on the grid based level of the CSTs. The validation/evaluation shows that the determination of structure variables on the individual tree level can only be carried out with limitations due to the specific stand characteristics (former coppice management). Particularly the horizontal crown dimensions and the amount of stems cannot be extracted with sufficient accuracy. The vertical dimensions could be determined with a good reliability, whereby a direct comparison of these results to existing studies is problematic due to the different forest structure characteristics.
The detection of the occurrence and the height of understory as well as the classification into deciduous and coniferous trees were performed with high accuracies. For these investigations the availability of the leaf-on/leaf-off data with a high point density was significant. The

determination of bio-physical parameters was possible to a certain extent; but for a sufficient validation an extension of the reference data is necessary.
We conclude that the direct, physically-based extraction of forest structure variables at the individual tree level has some limitations, particularly for the derivation of the tree/ stem locations and the horizontal crown dimensions. The presented approach based on CSTs provides robust and transferable information on the vertical canopy dimensions (e.g. amount of canopy layers, mean canopy height, mean height to canopy base, length of live canopy), understory characteristics and species distribution. However, the derivation of forest structure variables at the tree level is subject to larger errors due to omission and commission (as e.g. for stand density). Still, the information contained in the CSTs can easily be applied to improve the indirect/ empirical derivation of forest structure variables, depending on additional available in-situ data and the forestry expertise of users.

Acknowledgements

ESA is greatly acknowledged for funding the 3D-VegetationLab project, where this study is embedded in.

References

Adams, R., 1993. Radial decomposition of discs and spheres. *Graphical Models and Image Processing*, 55 (5), 325-332.

Antonarakis, A.S., Saatchi, S.S., Chazdon, R.L., and Moorcroft, P.R., 2011. Using Lidar and Radar measurements to constrain predictions of forest ecosystem structure and function. *Ecological Applications*, 21 (4), 1120-1137.

Betts, A.K., Ball, J.H., and McCaughey, J.H., 2001. Near-surface climate in the boreal forest. *Journal of Geophysical Research D: Atmospheres*, 106 (D24), 33529-33541.

Clark, P., and Evans, F. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35 (4), 445-453.

De Groot, R.S., Wilson, M.A. and Boumans, R.M.J, 2002. A typology for the classification, description and valuation of ecosystem functions, goods and services. *Ecological Economics*, 41 (3), 393-408.

Disney, M., Lewis, P., and Saich, P, 2006. 3D modeling of forest canopy structure for remote sensing simulations in the optical and microwave domains. *Remote Sensing of Environment*, 100 (1), 114-132.

Eugster, W., Zeyer, K., Zeeman, M., Michna, P., Zingg, A., Buchmann, N., and Emmenegger, L., 2007. Nitrous oxide net exchange in a beech dominated mixed forest in Switzerland measured with a quantum cascade laser spectrometer. *Biogeosciences Discussion*, 4 (2), 1167-1200.

Evans, J.S., and Hudak, A.T., 2007. A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments“, *IEEE Transactions on Geoscience and Remote Sensing*, 45 (4), 1029-1038.

Foody, G.M., 2010. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, 114 (10), 2271-2285.

Haara A., and Leskinen, P., 2009. The assessment of the uncertainty of updated stand-level inventory data. *Silva Fennica*, 43 (1), 87-112.

Hall, F.G., Bergen, K., Blair, J.B., Dubayah, R., Houghton, R., Hurtt, G., Kellndorfer, J., Lefsky, M., Ranson, J., Saatchi, S., Shugart, H.H., and Wickland, D., 2001. Characterizing 3D vegetation structure from space: Mission requirements. *Remote Sensing of Environment*, 115 (11), 2753-2775.

Hilker, T., Leeuwen, M. van, Coops, N.C., Wulder, M.A., Newnham, G.J., Jupp, D.L.B., and Culvenor, D.S., 2010. Comparing canopy metrics derived from terrestrial and airborne laser scanning in a Douglas-fir dominated forest stand. *Trees - Structure and Function*, 24 (5), 819-832.

Jones, T.G., Coops, N.C., and Sharma, T., 2012. Assessing the utility of LiDAR to differentiate among vegetation structural classes. *Remote Sensing Letters*, 3 (3), 231-238.

Jonsson, M., and Wardle, D.A., 2010. Structural equation modelling reveals plant-community drivers of carbon storage in boreal forest ecosystems. *Biology Letters*, 6 (1), 116-119.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., and Wu, J.C., 2012. An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, 4 (4), 950-974.

Kayes, L.J., and Tinker, D.B., 2012. Forest structure and regeneration following a mountain pine beetle epidemic in southeastern Wyoming. *Forest Ecology and Management*, 263, 57-66.

Kim, S., McGaughey, R.J., Andersen, H.-E., and Schreuder, G., 2009. Tree species differentiation using intensity data derived from leaf-on and leaf-off airborne laser scanner data. *Remote Sensing of Environment*, 113 (8), 1575-1586.

Korpela, I., Ørka, H.O., Maltamo, M., Tokola, T., and Hyyppä, J., 2010. Tree species classification using airborne LiDAR - effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, 44 (2), 319-339.

Leeuwen, M. and Nieuwenhuis, M., 2010. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, 129 (4), 749-770.

Lindenmayer, D.B., Franklin, J.F., and Fischer, J., 2006. General management principles and a checklist of strategies to guide forest biodiversity conservation. *Biological Conservation*, 131 (3), 433-445.

Liu, C., Frazier, P., and Kumar, L., 2007. Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107, 606-616.

Morsdorf, F., Nichol, C., Malthus, T., and Woodhouse, I.H., 2009. Assessing forest structural and physiological information content of multi-spectral LiDAR waveforms by radiative transfer modeling. *Remote Sensing of Environment*, 113 (10), 2152-2163.

Morsdorf, F., Kotz, B., Meier, E., Itten, K., and Allgower, B., 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. *Remote Sensing of Environment*, 104 (1), 50-61.

Mücke, W., Briese, C., and Hollaus, M., 2010. Terrain echo probability assignment based on full waveform airborne laser scanning observables. *Int. Archives of Photogrammetry and Remote Sensing*, XXXVIII/7A, pp. 157-162, Vienna, Austria.

Nadkarni, N.M., McIntosh, A.C.S., and Cushing, J.B., 2008. A framework to categorize forest structure concepts. *Forest Ecology and Management*, 256 (5), 872-882.

Purves, D., and Pacala, S., 2008. Predictive models of forest dynamics. *Science*, 320 (5882), 1452-1453.

RIEGL, 2012. Products. Airborne Scanning. Datasheets. < http://www.riegl.com>.

Roberts, J., Tesfamichael, S., Gebreslasie, M., Aardt, J. van, and Ahmed, F., 2007. Forest structural assessment using remote sensing technologies: an overview of the current state of the art. *Southern Hemisphere Forestry Journal*, 69 (3), 183-203.

Ross, A.N., 2011. Boundary-layer flow within and above a forest canopy of variable density. *Quarterly Journal of the Royal Meteorological Society*, DOI:10.1002/qj.989.

Shugart, H.H., Saatchi, S., and Hall, F.G., 2010. Importance of structure and its measurement in quantifying function of forest ecosystems. *Journal of Geophysical Research*, 115 (4), 1-16.

Sierra, C.A., Loescher, H.W., Harmon, M.E., Richardson, A.D., Hollinger, D.Y., and Perakis, S.S., 2009. Interannual variation of carbon fluxes from three contrasting evergreen forests: the role of forest dynamics and climate. *Ecology*, 90 (10), 2711-2723.

Strand, G., Dramstad, W., and Engan, G., 2002. The effect of field experience on the accuracy of identifying land cover types in aerial photographs. *International Journal of Applied Earth*, 4 (2), 137-146.

Van Calster, H., Baeten, L., Verheyen, K., De Keersmaeker, L., Dekeyser, S., Rogister, J.E., and Hermy, M., 2008. Diverging effects of overstorey conversion scenarios on the understorey vegetation in a former coppice-with-standards forest. *Forest Ecology and Management*, 256 (4), 519-528.

Vauhkonen, J., Tokola, T., Packalén, P., and Maltamo, M., 2009. Identification of scandinavian commercial species of individual trees from airborne laser scanning data using alpha shape metrics. *Forest Science*, 55 (1), 37-47.

Wagner, W., Hollaus, M., Briese, C., and Ducic, V., 2008. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. *International Journal of Remote Sensing*, 29 (5), 1433-1452.

Xue, B.-L., Kumagai, T., Iida, S., Nakai, T., Matsumoto, K., Komatsu, H., Otsuki, K., and Ohta, T., 2011. Influences of canopy structure and physiological traits on flux partitioning between understory and overstory in an eastern Siberian boreal larch forest. *Ecological Modelling*, 222 (8), 1479-1490.

Yang, R., and Friedl, M.A., 2003. Modeling the effects of three-dimensional vegetation structure on surface radiation and energy balance in boreal forests. *Journal of Geophysical Research D: Atmospheres*, 108 (16), GCP 10-1 - GCP 10-11.

Principal Curves for Tree Topology Retrieval from TLS Data

Anita Schilling[1], Anja Schmidt[1], Hans-Gerd Maas[1]

[1]Institute of Photogrammetry and Remote Sensing, Technische Universität Dresden,
anita.schilling@tu-dresden.de, anja.schmidt@tu-dresden.de,
hans-gerd.maas@tu-dresden.de

Paper Number: 042

Abstract

Reconstruction of a tree's topological and geometrical structure from terrestrial laser scanner point clouds is a fundamental step in order to gain insight into plant processes like water interception or light absorption. We propose to utilize principal curves as a novel approach to retrieve skeletal structures from TLS data sets. A principal curve is a polygonal line that traces the shape of a 3D point cloud by minimizing the expected squared distance to the given data set. Experiments are conducted to assess the feasibility and benefit of applying principal curves to single tree data sets. A previous segmentation of each tree into its phyto-elements is necessary and has been performed manually for simplicity. However, instead of introducing an arbitrary neighborhood structure like a voxel grid or graph, we propose to exploit the implicit neighborhood structure provided by each single TLS scan to facilitate data processing. Our results show that principal curves are an excellent tool for the retrieval of skeletal structures from point clouds and a fully automatic preprocessing should be aimed at in the future.

1. Introduction

Terrestrial laser scanners (TLS) have proven to be invaluable for capturing plant geometry data as 3D point cloud in a very efficient and precise way. However, the diversity in tree vegetation geometry and appearance makes it difficult to derive general constraints for tree reconstruction. Further complicating factors are technique-related effects during scanning like data gaps caused by occlusions, especially self-occlusions of branches in the tree crown, and artifacts due to wind movements. Moreover, the number of measured 3D points has a significant impact on the processing performance and the quality of results. In essence, processing scanned 3D point clouds with the objective to retrieve a skeletal representation of a tree's spatial structure is still a rather challenging task.

A common approach to deal with unorganized 3D point sets is the utilization of a voxel space representation of the data. (Gorte and Pfeifer 2004; Gorte and Winterhalder 2004) employed connected component labeling and mathematical morphology to carve out the skeleton of a tree from a voxel space representation. In (Gorte 2006), the procedure was improved further by incorporating Dijkstra's shortest path algorithm (Dijkstra 1959). A similar method was presented in (Gatziolis *et al*. 2010). A graph-reduction approach for skeleton extraction was introduced in (Bucksch 2011) mapping the point cloud onto an octree structure, which is then subjected to a set of fixed rules. However, the procedure tends to cause loops in the skeleton during graph processing and is computationally expensive. Recently, (Schilling *et al*. 2012) presented a method to retrieve a spatial tree representation utilizing Depth-First Search on the voxel representation of single tree point clouds. In general, voxel space approaches have the advantage of providing a fixed neighborhood structure that the raw 3D point cloud in most cases lacks. Nevertheless, methods based on voxel grids are usually hard to handle regarding the steering parameters and processing rules, and are computationally demanding.

In contrast, the raw 3D point set was used in (Livny *et al*. 2010) to create a weighted graph from which skeletal structures were retrieved by Dijkstra's algorithm. A similar approach was presented in (Côté *et al*. 2011), where intensity information was also taken into account, and

(Xu *et al*. 2007). In (Preuksakarn *et al*. 2010), the tree skeleton was retrieved by application of a point set contraction algorithm (Giannitrapani and Murino 1999). A probabilistic approach using general knowledge to guide a tree skeleton reconstruction process with iterative cylinder fitting was introduced in (Binney and Sukhatme 2009). In (Yan *et al*. 2009), an adjacency graph was built between point clusters previously established by repetitive k-means and fitting of minimum boundary cylinders. Furthermore, a method segmenting the range image of the point cloud based on local curvature estimates is detailed in (Dai *et al*. 2010). Subsequently, the skeleton is recovered from bins of clusters resulting from a previous region growing procedure. A similar approach utilizing principal curvature is presented in (Cheng *et al*. 2006). Since point clouds acquired by powerful TLS such as the Z+F Imager 5006i are rather large in size, computation of a nearest neighbor graph and application of graph algorithms becomes infeasible. Furthermore, the problem of occlusions, artifacts, and larger data gaps still remains.

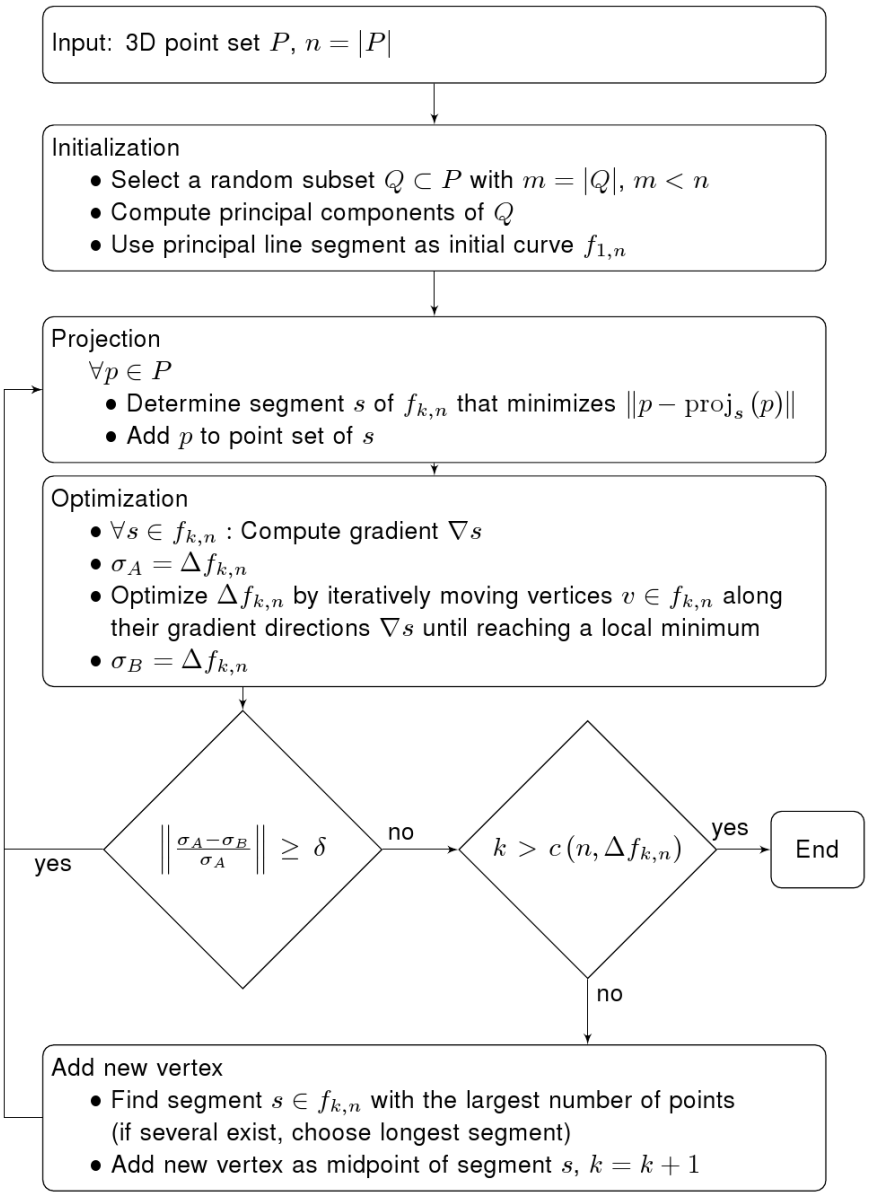
In this paper, we propose a novel approach to retrieve skeletal representations of trees from 3D point clouds on the basis of principal curves. A principal curve is a polygonal line that provides a compact summary of the point distribution of the 3D data and can be computed automatically. For this reason, principal curves are an obvious choice to trace point subsets representing tree branches; but to our knowledge, they have not been applied to TLS data before. Consequently, we present experiments that have been conducted in order to assess the feasibility of utilizing principal curves on TLS data.

The paper is organized as follows: First, principal curves and their computation scheme is introduced. Second, the study site and data preprocessing are detailed. Subsequently, a discussion of the results is presented. The paper closes with a summary of the conducted work in the conclusion.

2. Principal Curves

A 3D point set measured by a terrestrial laser scanner is a noisy point sampling of the geometry of a real-world object. If information on the object's geometry is available, its recovery from the 3D point set becomes feasible. In case of a 3D point set representing a cylindrical object, fitting a 3D center line by means of principal component analysis (e.g. (Jolliffe 2002)) is a trivial task. However, if the object resembles a tube, as is commonly observed at tree branches, recovering the underlying space curve from the 3D point set is far more difficult. If a model for the 3D curve cannot be determined in advance, methods based on Ransac (Fischler and Bolles 1981) cannot be applied because they require a known curve model to assess the error of the data points during computation.

In order to recover the underlying 3D curve from a 3D point set representing a tree branch or trunk, we propose the application of principal curves as defined in (Kégl *et al*. 2000): A principal curve is a polygonal line that traces the shape of a 3D point cloud by minimizing the expected squared distance to the given data set. In case of a 3D point set of a branch, it passes along the center of the branch. As detailed in (Kégl 1999; Kégl *et al*. 2000), the principal curve can be retrieved automatically from the 3D point set P with n denoting the number of data points. The first principal component of a subset of P is utilized as an initial line estimate. Successively, the algorithm alternates in an inner loop between a repartitioning of the data points according to their projections onto the curve, and an optimization computation of the curve vertices. After completing the inner loop, a new vertex is added to the curve and the inner loop is repeated. The computation terminates when a criterion $c(n, \Delta f_{k,n})$ is fulfilled, which incorporates the number of curve segments k and an error measure $\Delta f_{k,n}$ assessing the curve node positions in relation to their corresponding point subsets. The resulting polygonal line $f_{k,n}$ provides a compact description of the spatial distribution of the 3D point set P. Algorithm 1 provides an outline of the computation scheme.



Algorithm 1: Principal curves computation scheme based on (Kégl *et al*. 2000). δ denotes a predefined threshold of relative change.

3. Methods

3.1 Study Site and Data Processing

The study site comprises a plain birch stand (Betula pendula) on an area of ca. 1.3 ha (160 m × 80 m) near Wilmsdorf, Germany. The birch stand has been scanned in leafless condition using the TLS Z+F Imager 5006i with a field of view of 360° in horizontal, 310° in vertical direction and an angular resolution of 0.018°; for TLS specifications see (Zoller+Fröhlich 2009). For each independent scan, the data is provided similar to a range image. But here, each pixel of the image matrix represents a data tuple t in the form $t = (p_X, p_Y, p_Z, d, i)$, offering the 3D coordinates $(p_X, p_Y, p_Z) \in \mathbb{R}^3$ as well as the intensity value $i \in [0..1]$ in addition to each measured range $d \in \mathbb{R}^+$. If the measurement has failed at a particular pixel then an empty tuple $t_e = (0,0,0,0,0)$ has been stored. Arranging all of one scan's data in the same fashion as its range image makes each measured point readily accessible by its 2D coordinates in the image matrix. Hence, introducing an arbitrary neighborhood structure like a voxel grid or graph is not required.

The particular scan itself provides an implicit neighborhood structure, which can be directly exploited and facilitates general data processing as image processing methods can be applied to the 3D data as well. Another advantage is the possibility to visualize all data aspects of the scan within the image matrix, which provides a better and intuitive insight into the data set.

From the obtained scans, a set of 10 trees has been selected in $7.25\,m$ to $20.57\,m$ distance from the scanner position and exported as follows: All pixels of the original image matrix are checked and those, which hold 3D coordinates that are further away from the predefined tree position in the XY plane than $3\,m$, are replaced by empty tuples. Subsequently, the image matrix is clipped to the size of the 2D bounding box of the remaining non-empty tuples. For each tree, the result is a cutout of the original image matrix of the single-view scan, as shown in figure 1a.

Since the result of the principle curve computation is a single polygonal line, each tree has to be segmented into disjoint point sets denoting branches and the trunk. The result of this preprocessing step is a new image matrix, where each tuple $\hat{t} = (p_X, p_Y, p_Z, d, i, l)$ has an additional element $l \in \mathbb{N}^+$ denoting the point subset it belongs to. If the pixel contains an empty tuple, the label $l = 0$ is assigned, which denotes *background*.

a) b)

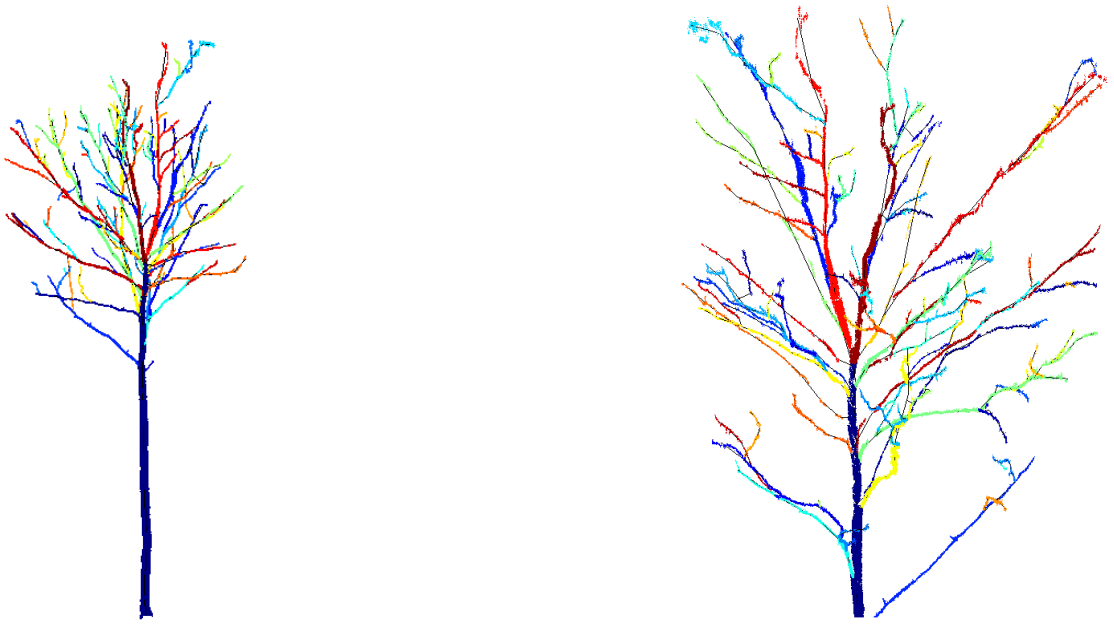
Figure 1: a) Range image cropped to bounding box. b) Manually enhanced label image, each color denotes a connected component determined by connected component labeling on the range image.

In order to assess the benefit of utilizing principle curves on TLS data, segmentation of phyto-elements has been performed manually on the basis of a previous connected component labeling step: For each of the obtained single tree images, the connected component labeling algorithm, as detailed in (Shapiro and Stockmann 2001), has been applied to the range values of the data tuples. Two neighboring pixels are defined to be *close* in 3D space and therefore connected pixels if the absolute difference of their range values does not exceed a predefined threshold ε. For the experiments, a threshold of $\varepsilon = 0.05\,m$ has been applied. Components representing a fragment of more than one branch due to a junction are split manually into separate components and a unique label l is assigned to each of them. Furthermore, the same label is assigned to disjoint components if they are representing the same branch according to

visual inspection. Very small components, which could not be assigned to a specific branch, have been eliminated. In that way, a unique label is determined for each separate phyto-element and propagated to the tuples of the pixels constituting the affected components as is demonstrated in figure 1b.

3.2 Experiment Setup

The 10 selected tree data sets have been prepared as explained in the previous section. Each input data set is treated as follows: For each unique label present in the current data set, a principal curve is computed for the corresponding 3D point subset. The point subset is obtained by selecting the 3D coordinates of all tuples where the label element matches the current label. Consequently, the result is a set of disjoint principal curves for each tree input data set. Linking the end points of the principal curve to attain a skeletal tree representation, as shown in figure 2b, could be performed automatically based on the segmented image (figure 1b), but has been conducted manually for simplicity. For the computation of the principal curves, an implementation in C++, which closely follows the prototype implementation provided in (Kégl 2000), has been utilized.



a) b) c)

Figure 2: Resulting skeletal representations of tree ID 8. a) Tree skeleton with 3D points colored according to label. b) Skeleton of tree. c) Magnified crown with 3D points and skeleton.

4. Results

Figure 2 gives an impression of the retrieved skeletal structures with an overlay of the corresponding point sets. The result of the principal curve computation for each single tree trunk is presented in table 1. Retrieval of the trunk curve is most time intensive due to its comparably high number of involved points. Clearly, the number of 3D points has a great impact on the time necessary to compute the principal curve. In addition, more iterations are generally required by the algorithm if the number of outliers or the outlier's distance error is large. The mean distance error denotes the average Euclidean distance of a 3D point to its corresponding node of the polygonal line; therefore, it should approximate the diameter of the trunk.

Table 1: Results of experiments for point subsets representing trunks.

Tree ID	Curve comp. time [s]	Mean distance error [m]	Number of points	Distance from scanner [m]
1	3.322	0.0667	82757	16.34
2	32.262	0.0755	357058	7.89
3	20.312	0.0721	332869	8.00
4	7.784	0.0671	140413	12.92
5	42.385	0.0699	361463	7.26
6	5.086	0.0904	120028	14.13
7	13.510	0.0710	184115	11.61
8	17.144	0.0837	196224	12.47
9	4.961	0.0798	84848	20.57
10	5.070	0.0796	88623	19.11

Computation results for branches are summarized for each tree in table 2. All branches are completed in less than $0.5\,s$ for point sets ranging from 20 to 13141 points in size. Again, the computation is quite robust to improperly segmented points, i.e. outliers to the targeted 3D point set representing the branch. For entire trees, the mean distance error denotes the average Euclidean distance of a 3D point to its corresponding node of the polygonal line over all branches in the crown. If the points have been segmented properly, the mean distance error of branch curves is less $1\,cm$ on average. Curve segments exhibiting a larger error in comparison to pre- and succeeding segments indicate the presence of outliers, e.g. due to a new branch forking off or improper segmentation. In those cases, the mean distance error is significantly higher, i.e. in the magnitude of several decimeters. The shape of the principal curve is governed strongly by the point distribution and drawn to more dense clusters of 3D points. For this reason, if the number of outlier points is small, the principal curve can still be retrieved correctly with only little noise influence. As indicated in figure 2c, data gaps in a branch can be bridged as well. Although winding branches in the scans, caused by wind during data capture, cannot be traced in all detail, the general branch shape can still be retrieved. In fact, the resulting straight, unwinding principal curve is more plausible in those specific cases.

Judged on visual inspection, the obtained tree structures trace the original spatial branching structure very well. Clearly, the quality of the results strongly depends on the prior segmentation of phyto-elements and correctly linking of computed principle curves.

Table 2: Results of experiment for each tree, excluding their trunk subsets.

Tree ID	Number of curves	Total Time [s]	Mean distance error [m]	Std.dev. of distance error [m]	Total number of points in crown	Number of points in largest subset
1	52	1.132	0.0009	0.0025	37717	11707
2	36	2.654	0.0007	0.0008	86473	12877
3	45	3.450	0.0024	0.0018	127748	9528
4	22	0.702	0.0084	0.0018	17468	2693
5	29	2.558	0.0014	0.0024	74141	11800
6	36	0.780	0.0074	0.0016	29003	6525
7	54	1.996	0.0071	0.0013	73956	10890
8	105	4.805	0.0042	0.0047	141898	13141
9	102	2.013	0.0081	0.0017	66036	4069
10	83	2.105	0.0074	0.0097	64142	7954

5. Conclusion

In this paper, we have proposed a novel approach to retrieve skeletal structures from TLS data on the basis of principal curves. A principal curve is a compact summary of a point cloud's distribution as polygonal line. To our knowledge, principal curves have not been applied to TLS data sets before. For this reason, we have conducted experiments on a set of 10 selected tree point clouds to assess the feasibility and benefit of employing principal curves on TLS data. Since each phyto-element of a tree has to be computed by a separate principal curve, a prior segmentation has been necessary. Rather than introducing an arbitrary neighborhood on the 3D point cloud in form of a graph or voxel grid, we have exploited the implicit neighborhood structure of the scan's range image. A connected component labeling algorithm has been performed on the range data in the image matrix. Subsequently, this initial segmentation has been improved by splitting or joining components thus that a unique label has been assigned to each branch or trunk. For each label, a principle curve has been computed and the resulting polygonal lines of a single tree have been connected to a tree graph.

Our results show that the principal curve computation is fast, robust to noise and can also deal with data gaps, which makes it well suited for application on 3D data sets. Furthermore, exploiting the implicit neighborhood structure of the range image matrix clearly facilitates general processing of 3D data from a single scan. In this way, well-known image processing methods can be adapted and applied to TLS data. Consequently, we strive to devise a fully automatic algorithm for retrieval of a complete and faithful tree model on the basis of principle curves in the future.

Acknowledgements

The work presented here was funded by DFG (Deutsche Forschungsgemeinschaft) via the project PAK331. The Z+F scanner was kindly provided by the Institute for Traffic Planning and Road Traffic (Prof. Dr. Lippold). We would especially like to thank our project partners, Prof. Dr. Wagner and Mr. Wollmerstädt, for their support in data acquisition.

References

Binney, J. and Sukhatme, G. S., 2009. 3d Tree Reconstruction from Laser Range Data. In: *Proceedings of the IEEE International Conference on Robotics and Automation*, Los Angeles, Universtiy of Southern California: 1321– 1326.

Bucksch, A., 2011. Revealing the skeleton from imperfect point clouds. PhD Thesis, TU Delft, Delft, Netherlands.

Cheng, Z., Zhang, X., and Fourcaud, T., 2007. Tree Skeleton Extraction from a Single Range Image. In: *Proceedings of the International Symposium on Plant Growth Modeling and Applications*, Beijing, CAS: 274-281, 2006.

Côté, J.-F., Fournier, R. A. and Egli, R., 2011. An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. *Environmental Modelling & Software*, 26, 761–777.

Dai, M., Li, H. and Zhang, X., 2010. Tree Modeling through Range Image Segmentation and 3D Shape Analysis. *Advances in Neural Network Research & Applications*, 67, 413–422.

Dijkstra, E.W., 1959. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1, 269–271.

Fischler, M.A. and Bolles, R.C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.

Gatziolis, D., Popescu, S., Sheridan, R., and Ku, N.-W., 2010. Evaluation of terrestrial LiDAR technology for the development of local tree volume equations. In: B. Koch, G. Kändler, C. Teguem (Eds.). *Proceedings of SilviLaser 2010 - The 10th International Conference on LiDAR Applications for Assessing Forest Ecosystems*, Freiburg, Germany: 197–205.

Giannitrapani, R. and Murino, V., 1999. Three-Dimensional Skeleton extraction by point set contraction. In: *Proceedings of the International Conference on Image Processing*, Udine, Udine University: 565-569.

Gorte, B. and Pfeifer, N., 2004. Structuring Laser-scanned trees using 3d mathematical morphology. *International Archives of Photogrammetry and Remote Sensing*, 35(B5), 929–933.

Gorte, B. and Winterhalder, D., 2004. Reconstruction of laser-scanned trees using filter operations in the 3D raster domain. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(8), 39–44.

Gorte, B., 2006. Skeletonization of Laser-Scanned Trees in the 3D Raster Domain. In: A. Abdul-Rahman, S. Zlatanova, V. Coors (Eds.). *Innovations in 3D Geo Information Systems*, Springer: 371–380.

Jolliffe, I. T., 2002. Principal Component Analysis. 2nd ed. Springer.

Kégl, B., 1999. Principal Curves: Learning, Design, and Applications. PhD Thesis, Concordia University, Montréal, Canada.

Kégl, B., 2000. Principal Curves. URL: http://www.iro.umontreal.ca/~kegl/research/pcurves/. Accessed: 22-Jun-2012.

Kégl, B., Krzyżak, A., Linder, T., and Zeger, K., 2000. Learning and design of principal curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3), 281–297.

Livny, Y., Yan, F., Olson, M., Chen, B., Zhang, H., and El-Sana, J., 2010. Automatic Reconstrution of Tree Skeletal Structures from Point Clouds. *ACM Transactions on Graphics*, 29(6), Article No. 151.

Preuksakarn, C., Boudon, F., Ferraro, P., Durand, J.B., Nikimmaa, E., and Godin, C., 2010. Reconstructing Plant Architecture from 3D Laser scanner data. In: T. de Jong, D. Da Silva (Eds.). *Proceedings of the 6th International Workshop on Functional-Structural Plant Models*, Davis, University of California: 16-18.

Schilling, A., Schmidt, A. and Maas, H.-G., 2012. Tree Topology Representation from TLS Point Clouds Using Depth-First Search in Voxel Space. *Photogrammetric Engineering and Remote Sensing*, 78(4), 383–392.

Shapiro, L.G. and Stockmann, G.C., 2001. Binary Image Analysis. In: *Computer Vision*, Prentice Hall: 69–75.

Xu, H., Gossett, N. and Chen, B., 2007. Knowledge and Heuristic-based Modeling of Laser-Scanned Trees. *ACM Transactions on Graphics*, 26(4), Article No. 19.

Yan, D.-M., Wintz, J., Mourrain, B., Wang, W., Boudon, F. and Godin, C., 2009. Efficient and robust reconstruction of botanical structure from laser scanned data points. In: *Proceedings of the 11th IEEE International Conference on Computer-Aided Design and Computer Graphics*, Hong Kong, University of Hong Kong: 572–576.

Zoller+Fröhlich, 2009. Datasheet Z+F Imager 5006i. URL:http://www.zf-laser.com/Datenblatt_IMAGER5006i_E.pdf , Accessed: 03-Aug-2011.

Identification of dead trees using small footprint full-waveform airborne laser scanning data

Werner Mücke[1], Markus Hollaus[1] & Norbert Pfeifer[1]

[1]Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Gußhausstraße 27-29, 1040 Vienna, Austria
[wm,mh,np]@ipf.tuwien.ac.at

Paper Number: SL2012-044

1. Introduction

Natura2000 is a European cross-border network for the monitoring of endangered and protected animal and plant species (European Commission 2012a). It is a part of the European Union's (EU) so-called Flora-Fauna-Habitat (FFH) directive, which was established in 1992 in order to conserve and protect wildlife and nature (European Commission 2012b). The FFH directive prescribes a regular monitoring cycle of six years for all Natura2000 regions, an area that in total currently covers approximately 778 000 km² (18%) of the EU's land surface. A number of ecologically relevant indicators for the monitoring were defined by the European Commission and expert consortia of the member states, which are to be used for the assessment of the condition of ecosystems. These indicators include the abundance of dead wood in forest ecosystems, which was identified as an important indicator for habitat condition (European Commission 2012c). It offers nutrition, shelter and housing for a number of animal species, and it is of major importance for carbon storage and forest productivity. The assessment of standing and fallen dead trees is therefore part of ecological monitoring and sustainable forest management. However, manual quantification of dead wood in forests is challenging, because extensive fieldwork in sometimes remote and rather inaccessible areas is required. Also, in terms of Natura2000 monitoring, the sheer extent of the areas makes it impossible to carry out the desired monitoring intervals using solely conventional ground-based mapping methods while still being cost-effective.
The EU-funded research project *Changehabitats2* (ChangeHabitats2 2012) aims to develop a monitoring methodology designed for the Natura2000 network, based mainly on remote sensing technologies. Methods of active and passive remote sensing shall be employed in order to support the conventional assessment process and, if possible, replace it partly. For this purpose airborne laser scanning (ALS) and hyper spectral (HS) image acquisition campaigns took place in selected areas in Germany and Hungary in 2011 and 2012. However, for this study we only involve the ALS data, as we want to test the ability of solely ALS to identify dead wood in forested areas. Also, we mainly concentrate on downed and standing dead trees that were partly or completely overgrown by surrounding trees. Their detection with HS data would anyway be difficult if not impossible, as HS imagery depicts the top most canopy and where it can penetrate it, it is highly influenced by object shadows. In several studies (Hyyppä et al. 2004; Næsset et al. 2004; Wagner et al. 2008; Wulder et al. 2012) it was shown that ALS data accurately depict the three-dimensional forest structure down to sub-canopy strata and the forest terrain, and can be employed to derive relevant metrics for quantitative and qualitative description of vegetation. Research has also shown that ALS data provide an excellent basis for the estimation of live and dead biomass in forests. For example, Vehmas et al. (2011) used ALS derived height metrics for the identification of canopy gaps with significant amounts of coarse woody debris (CWD). Bater et al. (2009) identified the coefficient of variation of ALS echo heights as the best variable for prediction of dead tree proportion. Kim et al. (2009) investigated the capability of ALS for the discrimination of live and dead standing tree biomass and found the echo intensity (often also referred to as echo amplitude) to be the most critical measure for

accurate estimation. Also Pesonen et al. (2008) derived ALS height density metrics for the prediction of standing and downed dead wood volume. All of the before mentioned examples have in common that they rely on statistical modeling and most often regression analysis for the identification of relevant ALS parameters, correlation analysis and prediction of downed and standing dead wood amounts. To our knowledge only one study exists that investigates the capability of ALS for direct mapping of fallen trees. Blanchard et al. (2011) used methods of object-based image analysis for the mapping of downed logs in gridded ALS data.
Considerable parts of this study are embedded in the on-going research in the *ChangeHabitats2* project, which has a lifetime till the end of 2014. Identification of downed and standing dead trees is a core product of the project. The presented approach for the detection of downed stems has some conceptual similarities to the method described by Blanchard et al. (2011), however we use different input parameters derived from full-waveform (FWF) ALS data. A stepwise process for the detection of fallen trees is proposed, which is based on the combination of point cloud filtering, morphological image processing and map algebra. As opposed to that, our efforts taken for the direct and automatic estimation of standing dead wood in the point cloud were inconclusive up to now. Nevertheless, we present the results of hitherto existing explorative point cloud analysis and give an outlook on future research on this topic within the project.
The following section 2 gives an introduction to the study area, the collected ground truth data and ALS data acquired for the project. Section 3 describes the workflow for the detection of fallen stems and the analysis of the ALS point cloud for the identification of the standing dead trees. All data analysis and processing was conducted using Mathworks Matlab and OPALS software packages (OPALS 2012). Section 4 shows and discusses the results, and section 5 concludes and gives an outlook on future related work.

2. Study area and data

The *Uckermark* is a landscape in north-eastern Germany, its biggest part being situated in the federal state of Brandenburg, the remaining parts in Mecklenburg-Vorpommern. The region is highly fertile, characterized by a number of bigger lakes (the so-called *Naturpark Uckermärkische Seen*) and by various rivers, which are often found together with alder-dominated riparian forests. Apart from that, beech forests, extensive pastures and wet meadows mostly dominate the *Uckermark*. The test site selected for this study is a beech stand (*fagus sylvatica*) with only few other individual species (*quercus robur, picea abies, fraxinus excelsior* and *carpinus betulus*). The structural characteristics of the stand vary from a hall-like appearance (old trees of 30 m and larger with only few to no understory) to successional younger parts (very dense or lots of shrub vegetation). There is only little variation in its relief and a few footpaths and driveways cross the area. Its extents are approximately 1200 x 900 m.
In 2011 and 2012 field-mapping campaigns for the purpose of Natura2000 mapping took place there. During the mapping a significant amount of dead trees (downed and standing) was found. The start- and endpoints of the downed stems (coarse woody debris) with a diameter bigger than 30 cm were measured with a differential GPS (dGPS), additionally their diameter and length were noted. The locations of the standing dead trees with a diameter at breast height (DBH) bigger than 30 cm and higher than 3 m were also measured with dGPS, also additionally their DBH and height were noted. Most of the standing dead trees were broken or decayed to a degree were crown-forming branches were no longer present. Terrestrial photographs were taken of all dead trees found in the area and stored together with the cardinal direction of the image acquisition. In total, 29 downed and 40 standing dead trees were mapped in the study site. Tree stumps and occasional piles of fine woody debris were not considered.
The ALS data were collected in spring 2011 (May 5th and 6th, leaf-on) and in early spring 2012 (March 22nd, leaf-off) using a Riegl LMS Q680i full-waveform system (RIEGL 2012a) mounted on a helicopter. A maximum deflection angle of ± 30°, a minimum lengthwise strip overlap of 50% and a pulse repetition rate of 400 kHz at an average flying height of 500 m a.g.l. resulted in

an average point density of 21.8 echoes/m² for the leaf-on and 16.9 echoes/m² for the leaf-off flight (considering all echoes). Echo (i.e. return) extraction from the raw full-waveform data was achieved by Gaussian decomposition (Wagner 2006) implemented in the software package RiPROCESS (RIEGL 2012b). A highly detailed digital terrain model (DTM, grid size 0.25 m) could be created from the last echoes of the leaf-off data using hierarchical robust filtering (Kraus and Pfeifer 1998) and subsequently the ALS point cloud was augmented with the normalized point heights (i.e. the height of the echoes above the terrain). The final point data set included the 3D coordinates of each echo, its amplitude and echo width (i.e. the width of the reflected pulse from full-waveform processing), the echo number, the total number of echoes per emitted pulse and the normalized point height.

3. Methods

3.1 Delineation of downed trees

The high-density point cloud available for this study provides good characterization of downed trees and made a direct identification of the stems possible. The proposed stepwise process (see Figure 1) begins with the selection of the ALS echoes with normalized heights less than 2 m and echo widths smaller than 4.5 ns. The 2 m height threshold is motivated by the fact that the biggest uprooted stems and their vertical root plates are commonly smaller than that. However, shrub vegetation would still be included in this subset and needs to be removed. For this the discriminative power of the echo width is exploited, which was investigated by recent research papers (Doneus et al. 2008; Hollaus et al. 2011; Mücke 2008). The echo width relates to small height variations of scattering elements within the footprint of the laser beam (Wagner 2005). It is therefore a measure of surface roughness and can be used to differentiate between rather smooth (in this case the bare forest ground or the downed stems, which appear also as a smooth surface in relation to the footprint size) and rather rough areas (here the shrub vegetation). In the proposed approach it is used in two ways: (1) directly as a threshold filter for the point cloud in the initial step of the process and (2) as a 0.5 x 0.5 m grid model of the echo width variance of the echoes in between 0.1 and 3 m distance to the DTM (referred to as EW_{var}) for classification purposes in the last step.

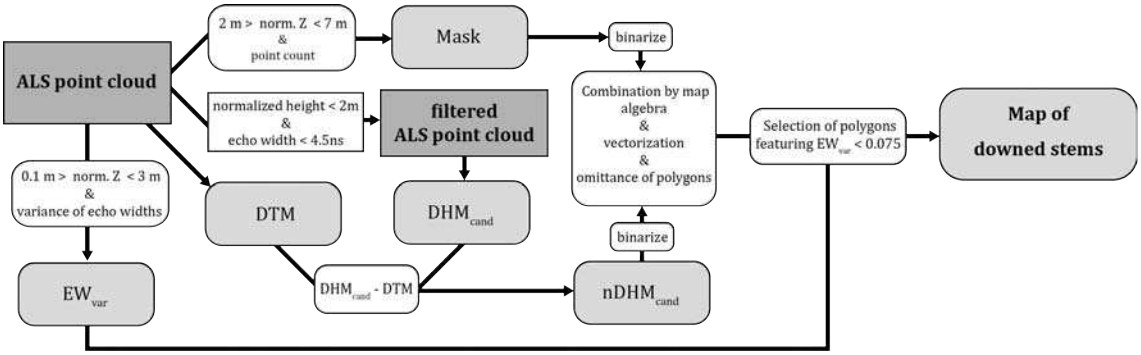


Figure 1: Schematic depiction of the identification workflow for downed stems.

The height and echo width threshold applied together on the point cloud significantly reduced the amount of data, while still including the echoes relevant for stem detection. A highly detailed digital height model is created from the remaining echoes by moving least squares interpolation (further referred to as DHM depicting the stem candidates, DHM_{cand}, grid size 0.25 m). It clearly shows the locations of the downed stems as elongated features (see Figure 2a). Then a difference model is calculated (normalized DHM_{cand}, $nDHM_{cand} = DHM_{cand} - DTM$), to which a height threshold of 10 cm is applied to separate the downed stems as off-terrain features. The resulting model is binarized (1 if $nDHM_{cand} > 0.1$ m else 0) and morphological closing is applied to connect stem areas that have been separated by the height thresholding and

to fill holes. The remaining image includes elongated areas representing the downed wood and spot-like features which are supposed to result from the stems of standing trees (see Figure 2b). To eliminate the spot-like areas we make use of the obvious fact that the stem of a standing tree continues above the initial 2 m height threshold and, due to the high point density, is represented by ALS echoes. A grid mask is created from the ALS echoes above 2 m and below 7 m including the point count per 1 x 1 m grid cell, which is then binarized (1 if the model exists, else 0) (referred to as $MASK_{bin}$). Using map algebra the grid models $nDHM_{cand}$ and $MASK_{bin}$ are combined and the resulting binary image is converted to vector format (ESRI shape), a process during which all polygons smaller than 3 m² are omitted because the chance for them being a stem is highly unlikely. The resulting vector map is spatially joined with the EW_{var} raster map and only the polygons featuring a small echo width variance (standard deviation of EW_{var} per polygon < 0.075) are selected. This final step further eliminates wrongly classified areas that are identified as near-ground vegetation due to higher surface roughness. The final vector map represents the outlines of the identified downed stems (see Figure 2c).

Visual comparison of the field-measured locations of the downed stems and the DHM_{cand} revealed that the positional accuracy of the stems mapped with dGPS was after all not good enough. Although every field-measured tree could be assigned a corresponding one in the DHM_{cand}, deviations of 1-3 m were observed, which made a direct automatic comparison with the automatic identification result (e.g. on pixel or object level) impossible. In fact, we found even more downed trees in the area through examination of the DHM_{cand}. To create a complete evaluation data set, the fallen stems that could be visually identified on the basis of the DHM_{cand} were manually digitized. The so created data set was then used for manual accuracy assessment of the automated method and concentrated only on the fact whether a stem was found (lengthwise overlap of reference and ALS-derived outlines > 75%), partly found (lengthwise overlap < 75%) or not found (no overlap).

3.2 Point cloud analysis for identification of standing dead trees

To investigate the representation of standing dead trees in the ALS point cloud, first sample trees had to be located in the study area. It was expected that the dGPS measured locations of the standing trees were prone to the same errors as the ones of the downed stems. Therefore, these locations were only used as initial positions for manually browsing the point cloud. From the data collected in the field, the height and diameter of the tree, as well as the information whether it still had a crown, was known and useful for identification of the trees. Twelve standing dead trees that distinguished themselves clearly (two of which still had a crown) could be selected. Their dGPS locations were subsequently corrected. Additionally, also twelve live trees were selected for comparison. The locations were used as centre points for a cylindrical extraction with a radius of 2.5 m of the ALS data. For the extracted 24 data sets explorative point cloud analysis was carried out. The different representations of live and dead standing trees concerning the point distribution (i.e. number of echoes in a certain height interval), the FWF attributes echo width and amplitude were investigated per sample tree. The investigations were carried out for the leaf-off and the leaf-on data set, likewise. For visualization purposes boxplots were created, showing the distribution of the tested echo attributes with respect to the corresponding normalized point height (see Figure 3 and Figure 4 **Error! Reference source not found.**).

4. Results and discussion

4.1 Delineation of downed trees

The results of the stepwise identification process are shown in Figure 2. The shading of the DHM_{cand} (Figure 2a) clearly shows a number of downed trees in this part of the study area, some of them have a vertical root plate. They are further highlighted by the binarized $nDHM_{cand}$

(elongated shapes in Figure 2b). This image also shows the spot-like areas that are supposed to represent stems of standing trees and remaining areas of near-ground vegetation, which were not removed entirely by the echo width filter applied to the point cloud. Figure 2c shows the final outlines of the automatically identified downed stems. In this part of the study area all but 2 stems were detected, 3 stems were only partly found. For the whole study area, out of a total of 193 manually digitized stems, 72 stems were fully detected (37.3%), 64 partly (33.2%) and 57 were not found (29.5%).

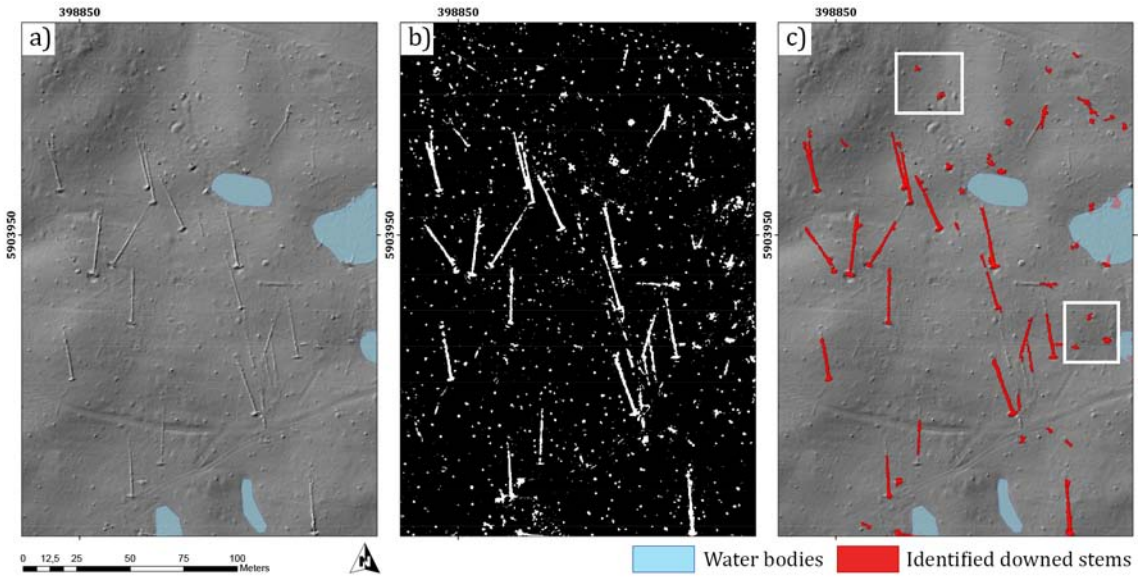


Figure 2: a) Shading of DHM_{cand} overlaid with manually digitized water bodies; b) binarized $nDHM_{cand}$; c) DHM_{cand} overlaid with identified downed stems (given coordinates are ETRS89 / UTM33 N).

The quality of the result is highly dependent on the DTM. On the one hand, the terrain modelling process (also referred to as filtering) has to be rather rigorous, as the initial echo selection is based on normalized point heights. Consequently, if the downed trees have a diameter smaller than 10 cm or are in an advanced state of decay, the height threshold applied to the $nDHM_{cand}$ excludes them. On the other hand, if the filtering is too rigorous, also small terrain bumps are filtered, which exhibit the same characteristics as downed stems (rather smooth areas in comparison to the laser beam and impenetrable, also often bigger than 3 m²). They therefore appear as spot-like polygons in the final result and are false positives (meaning polygons that were automatically classified as downed stems, but were not represented in the reference data), and cannot be excluded by FWF-ALS criteria. Some of the spot-like polygons in the final vector map (e.g. inside white rectangles in Figure 2c) were identified by the photographs taken in the field as (1) near-ground vegetation (shrubs and creepers), (2) piles of twigs or fine woody debris, and (3) vertical root plates of nearly fully decayed downed trees. We assume that (1) and (2) were too densely intertwined and hence featured comparable surface roughness to the downed stems. While case (1) is problematic, cases (2) and (3) are of course correct and a desirable result, but up to now our efforts to distinguish them were unsuccessful. Using additional geometric criteria, e.g. like perimeter-area ratio, might bring improvement, but evidently it would also delete objects that were identified correctly. Undoubtedly, this is the trade-off between completeness and correctness of the result, which has to be accepted using the proposed method.

4.2 Point cloud analysis for the identification of standing dead trees

For the explorative point cloud analysis, the ALS data were divided into subsets holding only first, last, single and all echoes, in order to investigate the influence of the different return types.

In concurrence with other existing studies we found the echo distribution and the echo amplitudes to be the strongest indicators for discrimination between standing live and dead trees. However, significant differences between the different return types were not observed. On the contrary, the more echoes were included in the analysis, the better the discriminatory power turned out to be. The boxplot given in Figure 3 shows that the echoes from dead trees are at both acquisition times more equally distributed than the ones for live trees. This is partly due to the high number of dead sample trees without a crown in our study area, but nevertheless it serves as an indicator for trees at this stage of decay.

John Arne
Deleted: F

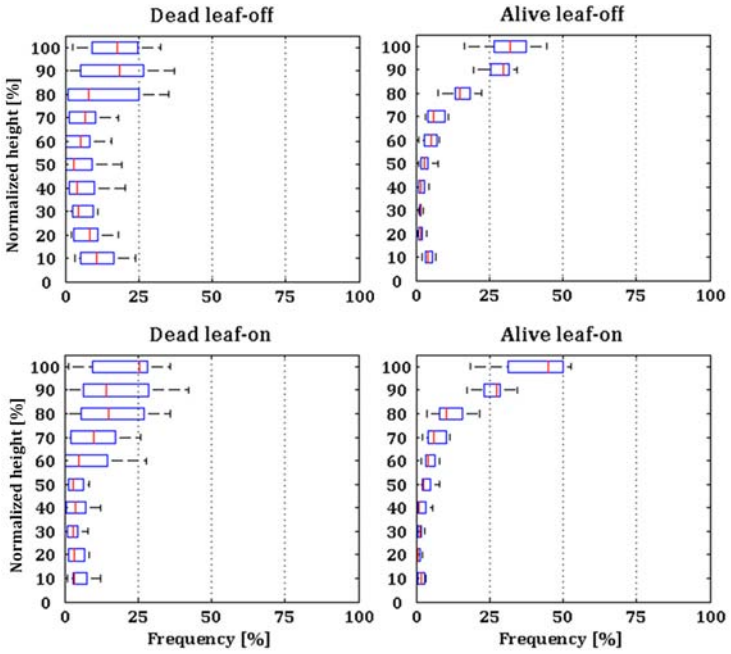


Figure 3: Boxplots showing the point distribution of all ALS echoes from the leaf-off and leaf-on data for the selected standing live and dead trees (red mark = median, blue box = 0.25 to 0.75 quantile, black line = data range without outliers).

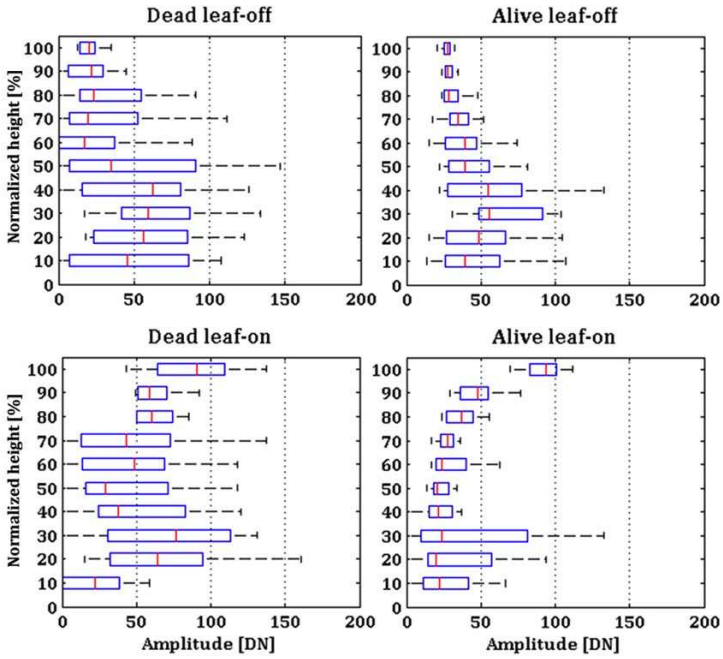


Figure 4: Boxplots showing the echo amplitudes of all ALS echoes from the leaf-off and leaf-on data for the selected standing live and dead trees (red mark = median, blue box = 0.25 to 0.75 quantile, black line = data range without outliers).

A relevant difference between leaf-off and leaf-on data could not be observed, although this is assumed to be a critical indicator for identifying standing dead trees that still have crown-forming branches, as it directly refers to vegetation phenology. Unfortunately, the subset of this tree type in our study area was too small to further elaborate on that matter.

In contrast, the echo amplitudes feature significant differences for leaf-off and leaf-on data, which can be seen in Figure 4**Error! Reference source not found.**. While again the distribution for the dead trees is rather equal for leaf-off and leaf-on, the trend for the live trees behaves contrarily. It shows significantly higher amplitudes in the top 30% of the echoes in the leaf-on than in the leaf-off, which is most likely caused by green foliage and its high reflectivity in the near infrared. This is also the strongest discriminator if employed for identification of dead trees during leaf-off season and live trees during leaf-on season using only the echoes in the top 30% of the tree heights. The high amplitudes in the leaf-on data set in the top most parts of the dead trees are most likely caused by branches of surrounding live trees, partly-overgrowing or growing into the dead tree.

John Arne
Deleted: F

5. Conclusion and outlook

This study investigates the capability of ALS for the identification of individual dead trees in forest ecosystems. The proposed workflow produces a vector map containing the outlines of the downed wood, which can further be used for clipping the ALS point cloud and selecting only echoes from downed trees. Also the vertical root plates are picked up in the data set of a density of approx. 20 echoes per m². Subsequently, cylinder fitting could be carried out in the 3D data to derive the dimensions of the stems. Based on the explorative point cloud analysis we envision a penetration index map on a grid basis for area wide identification of standing dead trees. It comprises the number of echoes in certain height intervals compared to all echoes and is therefore a measure of point distribution and penetration depth. Also, a map incorporating the ratio of amplitudes from the top 30% of all echoes (referring to the normalized echo height) from leaf-off and leaf-on ALS data can be conceived. Both maps will be subject in further studies and will be tested for their significance in identifying standing dead trees.

The information provided by the results of this study can be used to directly assess the state of ecosystems or, multi-temporal data acquisition assumed, highlight the ones where significant changes took place (so-called hot spots). Experts in FFH-mapping are then able to precisely select the Natura2000 habitats where they need to go and others, less affected or changed areas, can be treated with lower priority, thus improving the entire time- and cost-efficiency.

Acknowledgements

This study was partly funded by the *ChangeHabitats2* project (Marie Curie - FP7-PEOPLE-2009-IAPP - Grant Agreement Number 251234) and the *TransEcoNet* project implemented through the CENTRAL EUROPE Program co-financed by the ERDF.

References

Bater, Christopher W., Nicholas C. Coops, Sarah E. Gergel, Valerie LeMay, and Denis Collins. 2009. “Estimation of Standing Dead Tree Class Distributions in Northwest Coastal Forests Using Lidar Remote Sensing.” *Canadian Journal of Forest Research* 39 (6): 1080–1091.

Blanchard, Samuel D., Marek K. Jakubowski, and Maggi Kelly. 2011. “Object-Based Image Analysis of Downed Logs in Disturbed Forested Landscapes Using Lidar.” *Remote Sensing* 3 (11) (November 16): 2420–2439. doi:10.3390/rs3112420.

ChangeHabitats2, Project. 2012. “ChangeHabitats2 • For a Better Tomorrow.” http://www.changehabitats.eu/.

Doneus, Michael, Christian Briese, Martin Fera, and Martin Janner. 2008. “Archaeological Prospection of Forested Areas Using Full-waveform Airborne Laser Scanning.” *Journal of Archaeological Science* 35 (4) (April): 882–893. doi:10.1016/j.jas.2007.06.013.

European Commission. 2012a. “Natura 2000 Network - Environment - European Commission.” http://ec.europa.eu/environment/nature/natura2000/index_en.htm.

European Commission. 2012b. “The Habitats Directive - Environment - European Commission.” http://ec.europa.eu/environment/nature/legislation/habitatsdirective/index_en.htm.

European Commission. 2012c. “EU Biodiversity Indicators – SEBI 2010 - Environment - European Commission.” http://ec.europa.eu/environment/nature/knowledge/eu2010_indicators/index_en.htm.

Hollaus, Markus, Christoph Aubrecht, Bernhard Höfle, Klaus Steinnocher, and Wolfgang Wagner. 2011. “Roughness Mapping on Various Vertical Scales Based on Full-Waveform Airborne Laser Scanning Data.” *Remote Sensing* 3 (3) (March 4): 503–523. doi:10.3390/rs3030503.

Hyyppä Juha, Hannu Hyyppä, Paula Litkey, Xiaowei Yu, Henrik Haggrén, Petri Rönnholm, Ulla Pyysalo, Juho Pitkänen, and Matti Maltamo. 2004. “Algorithms and Methods of Airborne Laser-scanning for Forest Measurements.” *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (8/W2): 82–89.

Kim, Yunsuk, Zhiqiang Yang, Warren B. Cohen, Dirk Pflugmacher, Chris L. Lauver, and John L. Vankat. 2009. “Distinguishing Between Live and Dead Standing Tree Biomass on the North Rim of Grand Canyon National Park, USA Using Small-footprint Lidar Data.” *Remote Sensing of Environment* 113 (11) (November 16): 2499–2510. doi:10.1016/j.rse.2009.07.010.

Kraus, Karl, and Norbert Pfeifer. 1998. “Determination of Terrain Models in Wooded Areas with Airborne Laser Scanner Data.” *ISPRS Journal of Photogrammetry and Remote Sensing* 53 (4) (August): 193–203. doi:10.1016/S0924-2716(98)00009-4.

Mücke, Werner. 2008. “Analysis of Full-waveform Airborne Laser Scanning for the Improvement of DTM Generation”. Vienna: Vienna University of Technology.

Næsset, Erik, Terje Gobakken, Johan Holmgren, Hannu Hyyppä, Juha Hyyppä, Matti Maltamo, Mats Nilsson, Håkan Olsson, Åsa Persson, and Ulf Söderman. 2004. “Laser Scanning of Forest Resources: The Nordic Experience.” *Scandinavian Journal of Forest Research* 19 (6): 482–499. doi:10.1080/02827580410019553.

OPALS, Institute of Photogrammetry and Remote Sensing. 2012. “OPALS - Orientation and Processing of Airborne Laser Scanning Data.” http://www.ipf.tuwien.ac.at/opals/html/index.html.

Pesonen, Annukka, Matti Maltamo, Kalle Eerikäinen, and Petteri Packalèn. 2008. “Airborne Laser Scanning-based Prediction of Coarse Woody Debris Volumes in a Conservation Area.” *Forest Ecology and Management* 255 (8–9) (May 15): 3288–3296. doi:10.1016/j.foreco.2008.02.017.

RIEGL a, Laser Measurement GmbH. 2012. “RIEGL - Produktdetail.” http://www.riegl.com/nc/products/airborne-scanning/produktdetail/product/scanner/23/.

RIEGL b, Laser Measurement GmbH 2012, "RIEGL - Productdetail" http://www.riegl.com/index.php?id=232

Vehmas, Mikko, Petteri Packalén, Matti Maltamo, and Kalle Eerikäinen. 2011. “Using Airborne Laser Scanning Data for Detecting Canopy Gaps and Their Understory Type in Mature Boreal Forest.” *Annals of Forest Science* 68 (4): 825–835. doi:10.1007/s13595-011-0079-x.

Wagner, Wolfgang. 2005. “Physical Principles of Airborne Laser Scanning”. University Course.

Wagner, Wolfgang, Markus Hollaus, Christian Briese, and Vesna Ducic. 2008. “3D Vegetation Mapping Using Small-footprint Full-waveform Airborne Laser Scanners.” *International Journal of Remote Sensing* 29 (5): 1433–1452.

doi:10.1080/01431160701736398.

Wagner, Wolfgang, Andreas Ullrich, Vesna Ducic, Thomas Melzer, und Nick Studnicka. 2006. „Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner“. ISPRS Journal of Photogrammetry & Remote Sensing 60 (2): 100–112.

Wulder, Michael A., Joanne C. White, Ross F. Nelson, Erik Næsset, Hans Ole Ørka, Nicholas C. Coops, Thomas Hilker, Christopher W. Bater, and Terje Gobakken. 2012. “Lidar Sampling for Large-area Forest Characterization: A Review.” *Remote Sensing of Environment* 121 (0) (June): 196–209. doi:10.1016/j.rse.2012.02.001.

Testing a new vegetation structure retrieval algorithm from terrestrial lidar scanner data using 3D models

Mathias Disney[1,*1], Philip Lewis[2] & Pasi Raumonen[3]

[1]University College London, Geography, Gower Street, London, WC1E 6BT and NERC National Centre for Earth Observation mathias.disney@ucl.ac.uk
[2]University College London, Geography, Gower Street, London, WC1E 6BT and NERC National Centre for Earth Observation philip.lewis@ucl.ac.uk
[3] Tampere University of Technology, Department of Mathematics, P.O. Box 553, FI-33101 Tampere, Finland pasi.raumonen@tut.fi

Paper Number: SL2012-049

Abstract
Raumonen et al. (2011; 2012) have developed a new method for reconstructing topologically consistent tree architecture from terrestrial lidar scanner (TLS) point clouds. The method generates a cylinder model tree structure using a stepwise approach. The point cloud is covered with small sets that are along the tree surface. The cover sets are then segmented into branches that can be approximated with cylinders, checking connectivity at each stage. This approach avoids the need for restrictive assumptions about crown form, and automatically generates statistical properties of the reconstructed tree. The method has been shown to be effective for branch/trunk parts where point coverage is near complete, and the diameter of the part is large compared to the measurement accuracy. A difficulty of testing the method, and quantifying the impact of instrument and survey characteristics on the resulting structural information, is collecting sufficient accurate measurements from real trees under controlled conditions. Here, we overcome this by using a detailed 3D structural tree model and Monte Carlo ray tracing (MCRT). Thus 'tree' structure is known *a priori* and simulated TLS properties can be strictly controlled. We use a 3D model Scots pine tree, generated using a tree growth model and constructed of cylinders (although this could equally be done with facets or other geometric shapes), for which all positions and sizes are specified. TLS returns were simulated from three positions, using the same instrument characteristics as Raumonen et al. (2012) and 3D reconstruction was applied to the resulting point cloud. We show that the reconstruction works well on the simulated point cloud, with some constraints, reproducing the original branch length and volume with relative error < 2% for some cases. Accuracy is closely-linked to the size of the point subsets used in reconstruction. These results deonstrate the 3D simulation approach is ideally suited to testing reconstruction methods, which is an important part of demonstrating the utility of such methods.

1. Introduction

Vegetation structure is an important property that, in part, governs the transfer of gas, water and radiation at the land surface. Structure i.e. the size, shape and 3D arrangement of woody and leaf material, varies hugely within and across vegetation types and biomes, and is a difficult property to measure on anything other than

[1] Corresponding author.

plot-scale. Remote sensing measurements can be used to estimate structural properties indirectly. However such measurements typically require significant assumptions or simplifications regarding structure, and also tend to be a function of biochemical properties of the vegetation (Omasa et al., 2007). Lidar instruments provide much more direct estimates of structural properties, but in turn require different methods for interpreting the resulting data. Here, we focus on new methods for extracting 3D structural information from TLS point clouds.

Many studies have extracted 'canopy average' (spatially averaged) structural properties from TLS data e.g. leaf area index (LAI), gap probability, stem volume, basal area and above ground woody biomass (eg Omasa et al., 2007; Jupp et al., 2009; Yao et al., 2011; Seidel et al., 2012). Lovell et al. (2011) used TLS data to validate a simplified model of canopy scattering. New full-waveform TLS and systems with multiple wavelengths hold great promise for improved structural estimates (Jupp et al., 2009; Gaulton et al., 2011; Yao et al., 2011).

It has proved much harder to generate topologically consistent explicit 3D structural information from TLS data (as opposed to canopy averages). Reasons for this include the problem of occlusion within the point clouds and limited availability of highly-controlled TLS data for development and testing. But perhaps most importantly, testing 3D reconstruction methods requires tree structure to be known precisely *a priori*; this is very difficult in practice (Côté et al., 2009). However, if explicit 3D structure can be derived from TLS data rapidly and accurately, this would be extremely useful for applications including testing and benchmarking new models and parameter retrieval methods (Widlowski et al., 2008), simulating new sensors (Hancock et al., 2012) and driving radiation scattering models (Disney et al., 2006, 2011).

Growing interest in TLS systems has led to various approaches to 3D tree reconstruction. Hyyppä et al. (2001) used a segmentation–based method to retrieve crown stem volume. Pfeifer et al. (2004) and Pfeifer and Winterhalder (2004) used free-form curves to derive 3D structure from TLS data. Gorte and Pfeifer (2004) extended this approach to a more general tree morphological description. Bucksch and Lindenbergh (2008) used an octree structure to drive a skeletonisation approach. Binney and Sukhatme (2009) proposed a predictive model, working up in segments though the point cloud and using a simple sensor model to generate possible hypotheses for the next segment. Rutzinger et al. (2010) simplified TLS point cloud data into crown archetypes (cone, ellipse etc.), before applying a tree growth model to estimate within-crown tree branching structure. Côté et al. (2009) used a statistical method based on Verroust and Lazarus (2000) to reconstruct approximate trunk and branch topology. The approach was later extended to estimate 'stand average' trunk diameter and foliage area (Côté et al., 2011).

For all examples described above reconstructed canopy architecture was assessed either qualitatively, or using volume-average or crown-level comparisons. Côté et al. (2009) note that comparisons are necessarily limited by the requirement for a complete description of the structural attributes of the reference trees. This issue is exacerbated due to the difficulty of obtaining TLS data with well-characterised properties and minimal uncertainty (e.g. wind effects). However, these problems can be overcome by using an entirely simulated dataset for both the original reference 3D 'trees' and the TLS data. In this way 3D tree structure is known precisely *a priori*, and the characteristics of the TLS point clouds can be controlled. As long as 'sufficiently' realistic TLS point clouds can be simulated the test is then only of the performance of

the reconstruction method. Here, we present results of such an approach using 3D structural models and Monte Carlo ray tracing (MCRT) to test a new method of automated 3D tree reconstruction from TLS point clouds, which is both efficient and requires very few assumptions (Åkerblom, 2012; Åkerblom et al., 2012; Raumonen et al., 2011, 2012).

2. Methods

The method comprises three stages: i) MCRT simulation of TLS point clouds, using a 3D tree model as the 'truth' (known structural properties); ii) reconstruction of the tree structural properties (length, radius, volume of cylinders comprising the tree); and iii) comparison between original and reconstructed 3D structure.

2.1. Simulation

The 3D structural model used here represents a 30 yr old Scots pine tree, generated using an empirical growth model parameterised by species-dependent branching statistics in conjunction with specified external environmental conditions (Leersnijder, 1992). The specific tree model used here has been used in various studies (see Disney et al., 2006 fig. 4). TLS point clouds were simulated using the *librat* MCRT model. *librat*[2] is a MCRT simulation library that has been tested and validated against EO measurements and other models, and been used for a wide range of EO applications (Disney et al., 2006, 2011; Widlowski et al., 2008; Hancock et al. 2012). MCRT involves estimating canopy radiative transfer through stochastic sampling of possible photon 'ray' trajectories from source to sensor (forward) or sensor to source (reverse), and moderating photon attenuation at each ray interaction by specified material reflectance and transmittance properties. In lidar mode, the *librat* source and sensor are coincident, and 'photon' paths provide the distance to the first material interaction (for discrete return lidar systems) but also a distance-resolved signal for full-waveform simulations. Here, simulations were as for Raumonen et al. (2012) who used a Leica HDS6100 with a 650-690 nm TLS system (although the system is phase-based rather than discrete return the point cloud has the same properties) i.e. circular beam diameter at the exit of 3 mm and beam divergence of 0.22 mrad. The simulated lidar was situated at 1.5m height, at a radial distance of 20 m from the tree (at 0, 0, 0), at 45°, 0 and -45° in the xy plane i.e. xyz locations [14.1, 14.1, 1.5], [20, 0, 1.5] and [-14.1, 14.1, 1.5].

2.2. 3D Reconstruction

3D tree structure was reconstructed from the simulated point clouds using the method developed by Raumonen et al. (2011, 2012; Åkerblom, 2012) and applied in Åkerblom et al. (2012). The method reconstructs global 3D tree architecture step-by-step via an advancing collection of small local point cloud sets (surface patches). The stepwise approach does not require general *a priori* assumptions about tree structure and allows global reconstruction from a series of smaller, more computationally manageable independent subsets of local surface patches. The method results in a database of structural characteristics which, in addition to the topological representation of tree structure, can be used to describe trunk/branch diameter, length, location, and branch angular distributions and position (phyllotaxy). The method is scale-independent and can be used to reconstruct the cylindrical model accurately down

[2] *librat* is available as unsupported source code distribution from www.geog.ucl.ac.uk/~plewis/bpms/src/lib.

to the measurement accuracy of the laser scanning.

The method assumes the point cloud is a locally uniform sample of the branch and trunk surface of which contains 'sufficient' (within some tolerance) spatial detail of the branches. The method comprises six steps: i) generation of cover sets (see Figure 1, left); ii) derivation of the neighbour-relation and geometric characteristics of the cover sets; iii) tree components identified based on cover set characteristics; iv) segmentation of components into branches, recognising bifurcations by testing local connectivity; v) fit cylinders to small sub-regions of segments via least squares (see Figure 1, right); vi) fill gaps between cylinder segments. Reconstruction of the simulated point clouds were carried out for varying cover set diameters d_c of 1.0, 1.4, 1.8 and 2.5cm. The MATLAB © code processed 4.8M simulated points in 71 mins (1 cm diameter), down to 12 mins (2.5cm) on an Apple MacBook Pro with 2.8GHz CPU and 8GB of RAM.

Reconstruction accuracy was assessed by comparing the original 3D model (true) canopy structural characteristics (branch, trunk length and volume), with those derived via reconstruction, both at the tree level, as well as via cylinder size distributions.

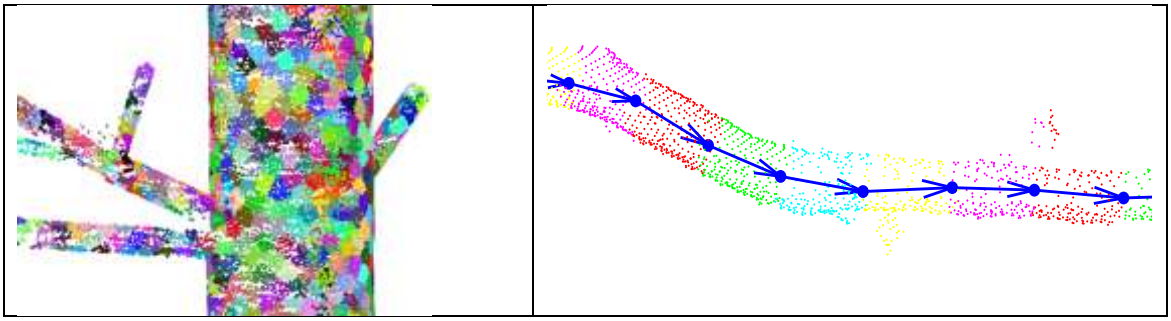


Figure 1 Left panel: cover sets (each of different colour) being formed on trunk and branch surfaces; right: a segment (branch) subdivided into regions (different colours) used for cylinder fitting, with arrows showing initially estimated axis directions.

3. Results

3.1. Tree-level comparison

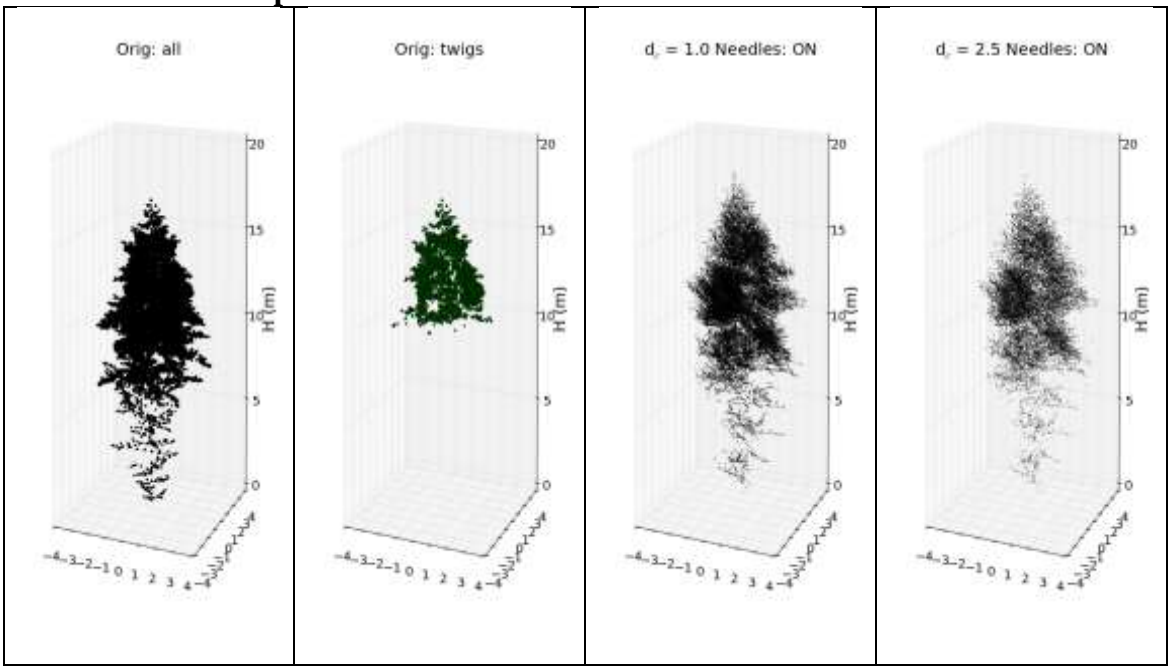


Figure 2 Left to right: cylinder vertices of original 3D model – all branches; twigs carrying needles only; reconstructed branch cylinder vertices with d_c = 1 and 2.5 cm respectively.

Table 1 Original and reconstructed branch, trunk and total length and volume.

		Reconstructed: varying cover set diameter, d_c (cm)			
	Original	**d_c = 1.0**	**d_c = 1.4**	**d_c = 1.8**	**d_c = 2.5**
Branch length (m)	2343	2093	2234	2266	2129
Branch volume (m^3)	0.320	0.298	0.329	0.364	0.522
Trunk length (m)	14.89	13.00	12.76	12.71	13.00
Trunk volume (m^3)	0.347	0.323	0.339	0.340	0.342
Total length (m)	2682	2106	2247	2279	2142
Total volume (m^3)	0.667	0.622	0.668	0.704	0.865
Height (m)	17.15	17.75	17.73	17.74	17.73
No. of cover sets		678925	417700	295809	206965

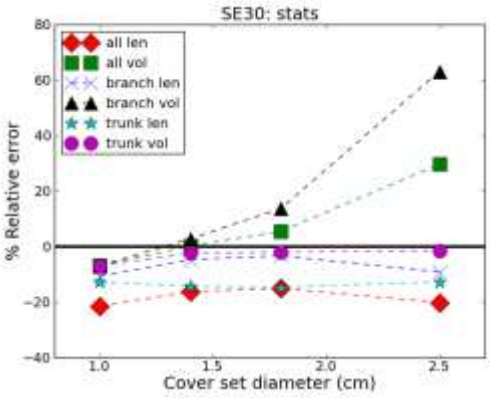


Figure 3 Comparison of original 3D tree model and reconstructed trees for varying cover set diameters.

The comparison of original 3D model cylinder vertices in Figure 2 shows good agreement with the reconstructed trees. The two right-hand panels show that as d_c increases from 1.0 and 2.5 cm some detail is lost. The comparison of cylinder length and volumes, along with total height, in Table 1 and Figure 3 quantify the differences between original and reconstructed 3D structure for increasing d_c i.e. lower resolution reconstructions. Figure 3 shows the variation between the original (O) and reconstructed (R) values, as the total % relative error (difference) i.e. 100 × (R-O)/O; where R are overestimates, the error will be positive. It can be seem that error in volume increases with d_c, while errors in length are flat, or slightly bell-shaped. The close agreement of cylinder size distributions are shown in Figure 5 and Figure 5 for d_c = 1.4 cm, the case with closest agreement of R_{volume} and R_{length} to the corresponding O values.

4. Discussion and conclusions

The results indicate that the 3D reconstruction method is able to reproduce the original 3D structure within a few percent. From Table 1, R_{height} is a very slight overestimate for all d_c, although the R_{branch} lengths are slight underestimates (3-10%). Underestimation of R_{trunk} is more significant (~13%). Figure 3 shows that as d_c increases, R_{volume} increases and so the error increases correspondingly. R_{volume} is very close to O_{volume} for d_c 1.0-1.8 cm, but a significant overestimate at d_c = 2.5 cm. Figure 4 and Figure 5 show that for all d_c, R_{radius} is closer to O_{radius} than R_{volume} to O_{volume}. For radii, there is consistent agreement across all values, with some offset. Volume is perhaps a better comparison of O v R, than length, as it accounts for the problem of single branches being reconstructed as multiple segments of shorter length. However, the highly skewed volume distribution (very many small, very few large) makes this somewhat tricky, although considering log bin counts helps. For both O and R there are

a very large number of cylinders with small volumes (95% of and 93% < 30cm^3 respectively), with the remaining cylinders spread over several orders of magnitude in volume, with a very few comparatively very large ones. This accounts for the much greater scatter at low numbers seen in the right panel of Figure 5.

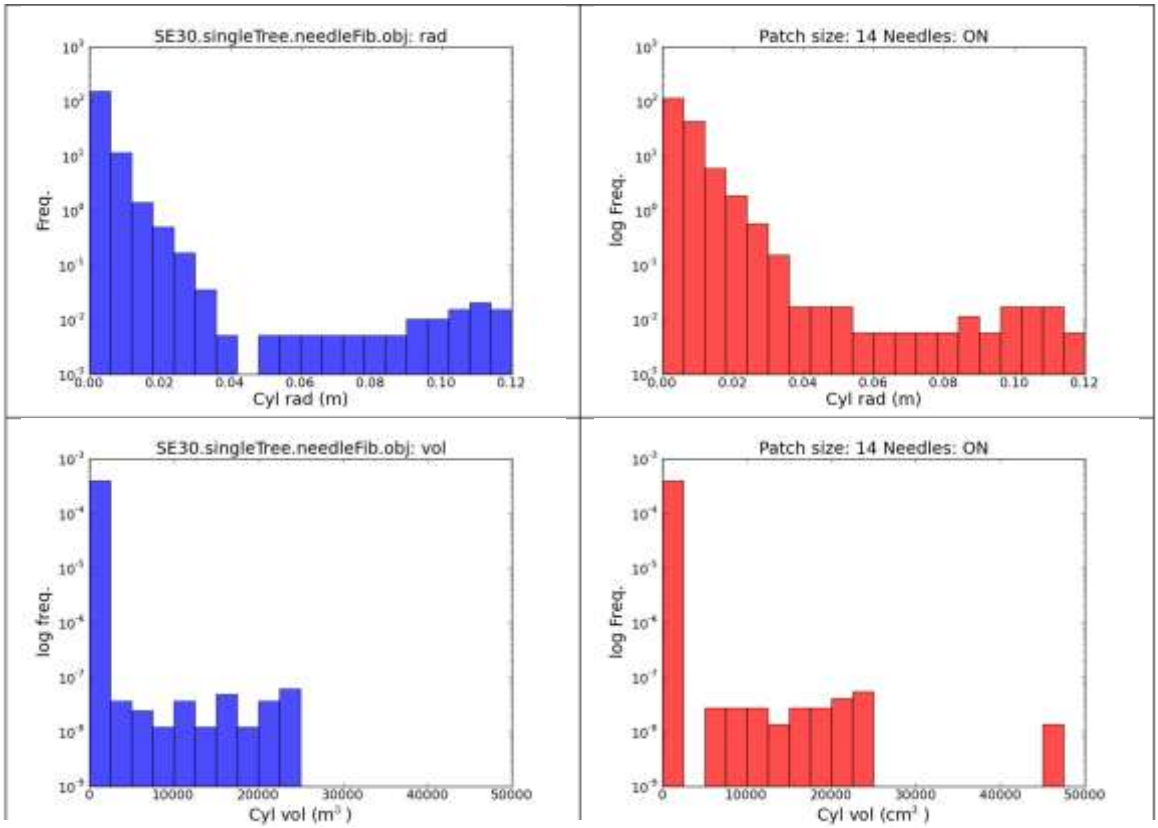


Figure 4 Top row: normalised histogram of original 3D cylinder radii (blue) and reconstructed radii (red) for d_c = 1.4 cm. Bottom row: as above, but for cylinder length.

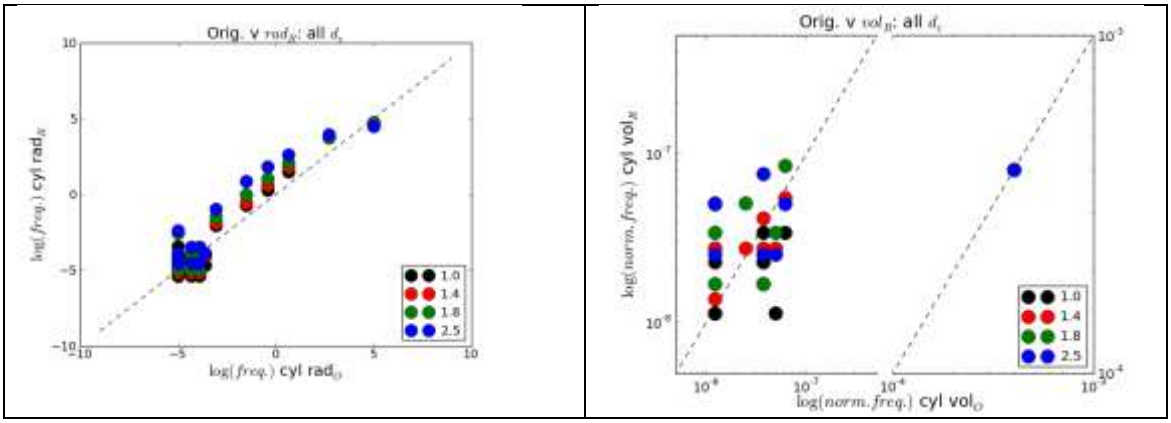


Figure 5 Scatter of log of normalised frequency distributions of original v reconstructed cylinder radii (left) and volume (right – note change of scale), for all d_c (1:1 line shown as dotted).

The underestimation of R_{trunk} in Table 1 indicates that automatic classification of the trunk components is not working perfectly. Identifying issues of this sort is precisely why the approach taken here is ideal for testing the reconstruction. Additional results (not shown) for reconstructions with needles removed indicate that needles (unsurprisingly) have a significant impact, increasing R volume; the needles make the

branches appear thicker in the TLS measurements. Figure 3 in particular indicates that there is an optimum value of d_c for reconstruction, below which there is a small decrease in accuracy, but above which accuracy declines rapidly. In this case, smaller branches close together cannot be separated and/or their bases may be poorly resolved: even if they are recognised as separate branches, part of their base will be in the parent segment, because a single cover set is larger than the branch size, so including both parent and child parts. Branches much smaller than d_c may not be recognised at all, and included in their parent branches. This will tend to cause such parent branches to be much thicker than they ought to be. This is also evident in the branch volume distributions, where peak volume of branches is shifted into larger branches with increasing d_c. For applications such as biomass estimation however, a slight underestimation of smaller branches will have little impact, particularly if the total volume is retrieved accurately, as suggested here.

In conclusion, we have used 3D model trees, combined with MCRT simulation to generate TLS point clouds with precisely known tree structure and TLS instrument characteristics. We have demonstrated the application of a new 3D tree structure reconstruction method can successfully reconstruct the original 3D tree information very well, with certain constraints, particularly the choice of d_c, the size of the cover set. Results are presented for a single (isolated) tree only, but the same approach can be used to examine the obscuration effect in a stand of trees. We are currently carrying out simulations exploring multiple trees of different size and structure, both singly and in stands.

3D reconstruction from TLS holds much promise for rapid, repeatable measurements of tree structure for a range of applications, but the difficulty has always been quantifying how good the resulting reconstructions are, and how sensitive the methods are to particular instrument characteristics and experimental limitations. We show that the approach of using known structure and controllable instrument properties is ideally suited to precise, quantifiable testing of TLS reconstruction methods.

References

Åkerblom, M. 2012. *Quantitative tree modelling from laser scanning data*. M.Sc. thesis, Tampere University of Technology.

Åkerblom, M., Raumonen, P. and Kaasalainen, M. 2012. Comprehensive quantitative tree models from TLS data. *In proc. IEEE International Geoscience and Remote Sensing Society Symposium*, Hamburg, Germany, 23-27 July 2012.

Binney, J. and Sukhatme, G. S. 2009. 3D tree reconstruction from laser range data, in *Proc. Robotics and Automation, ICRA '09 IEEE International Conference*, 12-17 May 2009, 1321-1326.

Bucksch, A. and Lindenbergh, R. 2008. CAMPINO - A skeletonization method for point cloud processing, *ISPRS Journal of Photogrammetry & Remote Sensing*, 63, 115-127.

Hyyppä, J., Kelle, O., Lehikoinen, M. and Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 969-975.

Disney, M. I., Lewis, P., Saich, P. 2006. 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains, *Remote Sensing of Environment*, 100(1), 114-132.

Disney, M. I., Lewis, P., Gomez-Dans, J., Roy, D., Wooster, M. and Lajas, D. 2011. 3D radiative transfer modelling of fire impacts on a two-layer savanna system, *Remote Sensing of Environment*, 115, 1866-1881.

Gaulton, R., Danson, F. M., Pearson, G., Lewis, P. E. and Disney, M. I. 2010. The Salford Advanced Laser Canopy Analyser (SALCA): A multispectral full waveform LiDAR for improved vegetation characterisation. In *Proc. Rem. Sens. and Photog. Soc. Conference*, Burlington House, London, 5th May 2010.

Gorte, B. and Pfeifer N. 2004. Structuring laser-scanned trees using 3D mathematical morphology. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXV, B5, pp. 929-933.

Hancock, S., Lewis, P., Foster, M., Disney, M. I., and Muller, J. -P. 2012. Measuring tree height over topography and understory vegetation with dual wavelength lidar: a simulation study. *Agricultural and Forest Meteorology*, 161, 123-133.

Jupp, D. L. B., Culvenor, D. S., Lovell, J. L., Newnham, G. J., Strahler, A. H. and Woodcock, C. E. 2009. Estimating forest LAI profiles and structural parameters using a ground-based laser called Echidna ®. *Tree Physiology*, 29(2), 171-181.

Leersnijder, R. P. 1992. *PINOGRAM: A pine growth area model*. WAU dissertation 1499, Wageningen Agricultural University, The Netherlands.

Lovell, J. L., Haverd, V., Jupp, D. L. B. and Newnham, G. J. 2011. The Canopy Semi-analytic Pgap And Radiative Transfer (CanSPART) model: Validation using ground based lidar. *Agricultural and Forest Meteorology*, 158–159, 1–12

Omasa, K., Hosoi, F. and Konishi. A. 2007. 3D lidar imaging for detecting and understanding plant responses and canopy structure, *Journal of Experimental Botany*, 58(4), 881–898.

Pesci, A., Teza, G. abd Bonali. E. 2011. Terrestrial Laser Scanner Resolution: Numerical Simulations and Experiments on Spatial Sampling Optimization. *Remote Sensing*, 3, 167-184, doi:10.3390/rs3010167.

Pfeifer, N., Gorte, B., Winterhalder D. 2004. Automatic reconstruction of single trees from terrestrial laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXV, B5, 114-119.

Pfeifer, N. and Winterhalder D. 2004. Modelling of Tree Cross Sections from Terrestrial Laser-Scanning Data with Free-Form Curves. *International Archives of Photogrammetry, remote sensing and spatial information science*s, Vol XXXVI, part 8/W2, 76-81.

Raumonen, P., Kaasalainen, S., Kaasalainen, M., and Kaartinen, H. 2011. Approximation of volume and branch size distribution of trees from laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(5/W12).

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., Lewis, P. 2012, submitted. Comprehensive quantitative tree models from terrestrial laser scanner data.

Rutzinger, M., Pratihast, A. K., Oude Elberink, S. and Vosselman, G. 2010. Detection and modelling of 3D trees from mobile laser scanning data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVIII, Part 5 Commission V Symposium, Newcastle upon Tyne, UK.

Seidel, D., Fleck, S. and Leuschner, C. 2012. Analyzing forest canopies with ground-based laser scanning: A comparison with hemispherical photography. *Agricultural and Forest Meteorology*, 154-155, 1-8.

Van der Zande, D., Stuckens, J., Verstraeten, W. W., Mereu, S., Muysf B. and Coppin, P. 2011. 3D modeling of light interception in heterogeneous forest canopies using ground-based LiDAR data. *International Journal of Applied Earth Observation and Geoinformation*, 13, 792–800.

Verroust, A., & Lazarus, F. 2000. Extracting skeletal curves from 3D scattered data. *The Visual Computer*, 16, 15−25.

Widlowski, J. -L., Robustelli, M., Disney, M. I., J.-P. Gastellu-Etchegorry, T. Lavergne, P. Lewis, P. R. J. North, B. Pinty, R. Thompson and M. M. Verstraete 2008. The RAMI Online Model Checker (ROMC): A web-based benchmarking facility for canopy reflectance models. *Remote Sensing of Environment*, 112(3), 1144-1150.

Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D., Lovell, J., Culvenor, D, Newnham, G., Ni-Meister, W., Schaaf, C., Woodcock, C., Wang, J., Li, X. and Strahler, A. 2011. Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. *Remote Sensing of Environment*, 115(11), 1144-1150.

Preliminary investigation of the accuracy of global canopy height products over Canada using airborne Lidar data

Douglas K. Bolton[1], Nicholas C. Coops[1], Michael A. Wulder[2]

[1]Department of Forest Resources Management, Faculty of Forestry, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada, dbolt@interchange.ubc.ca, nicholas.coops@ubc.ca

[2]Canadian Forest Service, Natural Resources Canada, Victoria, British Columbia, V8Z 1M5, Canada, mwulder@nrcan.gc.ca

Abstract Submission #SL2012-062

Abstract

Carbon storage in forest aboveground biomass is a critical, yet difficult, component of the global carbon cycle to estimate. Recent attempts to measure forest height, an indicator of site carbon storage, over large areas have relied on empirical relationships between globally available spatial layers and field measured heights. At the global level, moderate spatial resolution optical imagery (pixels > 100m) often can provide the spectral and spatial coverage to extrapolate canopy heights, while estimates of canopy heights themselves can be derived from the Geoscience Laser Altimeter System (GLAS) aboard the Ice, Cloud and land Elevation Satellite (ICESat). In this paper, we investigate the nature of differences between two published ICESat-derived products that estimate canopy height globally by comparing these products to measurements of height from small-footprint airborne Light Detection and Ranging (Lidar) data across Canada's forested ecozones. The first map of forest height, produced by Lefsky (2010), extrapolates GLAS derived heights using spectral data from the MODerate resolution Imaging Spectrometer (MODIS), while the second map, produced by Simard *et al.* (2011), extrapolates GLAS derived heights using climatic and additional ancillary variables. In 2010, the Canadian Forest Service (CFS) collected transects totaling over 25,000 km of small-footprint discrete return LiDAR data across boreal Canada, providing an opportunity to explore the patterns of differences. Results suggest that Lidar derived stand heights (95th height percentile) are more strongly correlated to forest height estimates from Simard *et al.* (2011, $R^2 = 0.49$) than to those from Lefsky (2010, $R^2 = 0.14$) at the ecodistrict level in Canada. Lefsky (2010) predicted taller canopy heights in most Boreal Cordillera ecodistricts than the small-footprint LiDAR while Simard *et al.* (2011) did not capture the spatial variations in height within the Taiga Plains. Discrepancies between the published heights maps in the Boreal Cordillera, Pacific Maritime and the Montane Cordillera reveal the influence that processing and extrapolation techniques can have on height estimates. Both heights maps appeared less able to predict height over sloped terrain. Investigating the relationship between small-footprint Lidar data and published forest height products can identify approaches that lead to more accurate estimates of aboveground biomass and can help determine why discrepancies exists in forest height estimates between various approaches, data, modeling, and underlying environmental conditions.

1 Introduction

Forests are an integral component of the global carbon (C) cycle, with terrestrial C uptake estimated at 1.0 ± 0.6 Gt yr^{-1} during the 1990's and 0.9 ± 0.6 Gt yr^{-1} between 2000-2005 (Denman *et al.* 2007). Projected changes in forest productivity (Boisvenue and Running 2006), disturbance regimes (Flannigan *et al.* 2005) and land conversion are expected to have a significant impact on carbon sequestration and storage in forest ecosystems (Le Quéré *et al.* 2009). As forests currently represent a large sink of anthropogenic carbon dioxide emissions (Denman *et al.* 2007), modeling changes in forest C dynamics is a critical component to forecasting future climate (Houghton *et al.* 2009).

Modeling forest C dynamics into the future requires an accurate characterization of current carbon storage in forest ecosystems. As changes in carbon storage will not be uniform globally an understanding of the spatial patterns of above-ground biomass (AGB) in forests is necessary to accurately parameterize carbon models (Houghton *et al.* 2009). The cost and time constraints of field-based measurements prevent the characterization of AGB over large areas through field studies alone (Gillis *et al.* 2005; Wulder *et al.* 2007; Masek *et al.* 2011). While capable of monitoring vegetation over continental to global scales, optical remote sensing is not well suited for capturing the vertical structure of vegetation, limiting its utility for characterizing AGB (Goetz and Dubayah 2011). Light Detection and Ranging (Lidar), on the other hand, can measure vertical structure of vegetation (Lefsky *et al.* 1999; Dubayah and Drake 2000), but cannot achieve the continental and global coverage of optical sensors.

The Geoscience Laser Altimeter System (GLAS) aboard the Ice, Cloud and land Elevation Satellite (ICESat) collected waveform Lidar data globally from 2003 to 2009. GLAS laser footprints are ~65 m in diameter and separated by 172 m along track with an 8- or 91-day repeat cycle (Zwally *et al.* 2002), providing a sample of forest structure over the globe during the 7 year mission. Two recent programs to provide wall-to-wall estimates of canopy height, a key indicator of AGB, have relied on extrapolating GLAS derived canopy heights through empirical relationships with globally available spatial layers (Lefsky 2010; Simard *et al.* 2011). Lefsky (2010) extrapolated GLAS derived heights using segmented spectral data and land cover classifications from the Moderate Resolution Imaging Spectrometer (MODIS), with segments ranging in size from 1 to 900 pixels (0.25-225 km^2). Simard *et al.* (2011) extrapolated GLAS derived heights using climate, elevation, MODIS tree cover and protection status layers at a consistent spatial resolution of 1-km. As GLAS derived height products can complement field-based inventories and provide valuable information on carbon storage in forests, the accuracy of these products and the errors they will introduce to carbon models must be further assessed and better understood.

In this analysis, we compare and contrast the GLAS derived height maps produced by Lefsky (2010) and Simard *et al.* (2011) over the forested ecozones of Canada. We compare the height maps against airborne Lidar data collected by the Canadian Forest Service (CFS) in 2010 in order to investigate map accuracy and gain an improved understanding of the conditions that lead to agreement and disagreement between the GLAS derived products. The findings are intended to inform the interpretations of information generated from these height products as well as providing insights for the generation of future large area height products with lidar, models, and remotely sensed data.

2 Methods

2.1 Study Area

Canada's forests occupy an area of approximately 400 million hectares that represents over 53% of the land area of Canada (Natural Resources Canada 2011) and forms ~10% of the world's total

forested area. Canadian forests occupy a wide range of climatic, topographic and soil conditions, ranging from highly productive temperate rainforests forests in British Columbia to smaller trees in slower growing forests in the more northern boreal (Ecological Stratitification Working Group 1995). The ability to characterize spatial variations in forest height across this range of conditions is critical to accurate predict current carbon storage in AGB. Figure 1 displays the forested ecozones of Canada, as defined by the Ecological Stratification Working Group (1995). The ecozones are further divided into ecoregions and ecodistricts, with ecodistricts representing the finest scale division of geomorphology, soil, vegetation and climate characteristics (Ecological Stratitification Working Group 1995). As we are interested in the ability of GLAS derived height products to capture spatial variations in canopy height, we compare and contrast height estimates at the ecodistrict and ecozone level across Canada.

2.2 Data

The global GLAS derived height maps produced by Lefsky (2010) and Simard *et al.* (2011) are both freely available online at a spatial resolution of 500 m and 1 km, respectively. Both height maps were averaged from the pixel to the ecodistrict level across Canada. Ecodistricts containing fewer than 1,000 ha of height estimates for each map (10 pixels for Simard *et al.*, 40 pixels for Lefsky), were removed.

To investigate map accuracy and the conditions that lead to agreement and disagreement between these heights products, several additional data sources were acquired. In the summer of 2010, 34 transects of small-footprint airborne Lidar data were collected by the CFS, totaling nearly 25,000 km in length and spanning from Newfoundland in the east to the Yukon in the west (Figure 1). The data were collected between the altitudes of 450 and 1900 m using an Optech ALTM 3100 discrete return sensor, with a pulse repetition frequency of 70 kHz and a fixed scan angle of 15° for the majority of transects (Wulder *et al. in press*). The transects were partitioned into 25 by 25 m plots and a suite of forest relevant metrics was generated for each plot using FUSION, a freely available software package for Lidar data processing developed by the USDA Forest Service (McGaughey, 2012). By summarizing the distribution and density of Lidar returns, these metrics provide a set of variables describing the three-dimensional structure of each plot. Height percentiles, which describe the cumulative distribution of Lidar returns, are strong predictors of forest attributes such as mean or dominant tree height. As physical or atmospheric anomalies can influence the 99th height percentile or maximum return height (Magnussen and Boudewyn 1998; Kane *et al.* 2010), the 95th height percentile was selected as an estimate of canopy height to compare against the GLAS derived height products in this analysis. The CFS assigned each Lidar plot with a land cover class, according to the 25-m resolution Earth Observation for Sustainable Development of Forests (EOSD) land cover classification (Wulder *et al.* 2008). Lidar metrics for all plots classified as forest by the EOSD were selected and averaged to the ecodistrict level for comparison against the GLAS derived height products.

As the accuracy of GLAS derived heights has been shown to be less reliable on sloped terrain (Duncanson *et al.* 2010; Xing *et al.* 2010; Pang *et al.* 2011), we assessed the effect of terrain on height estimates using data from the Shuttle Radar Topography Mission (SRTM) digital elevation model, which was produced near-globally at 90 m spatial resolution (Rabus *et al.* 2003). Terrain roughness was calculated for each ecodistrict as the standard deviation of all 90m SRTM pixels within the ecodistrict.

2.3 Comparison of GLAS derived height products

To investigate the agreement between the Lefsky (2010) and Simard *et al.* (2011) height products, the ecodistrict average heights were grouped according to ecozone and displayed as boxplots. Large

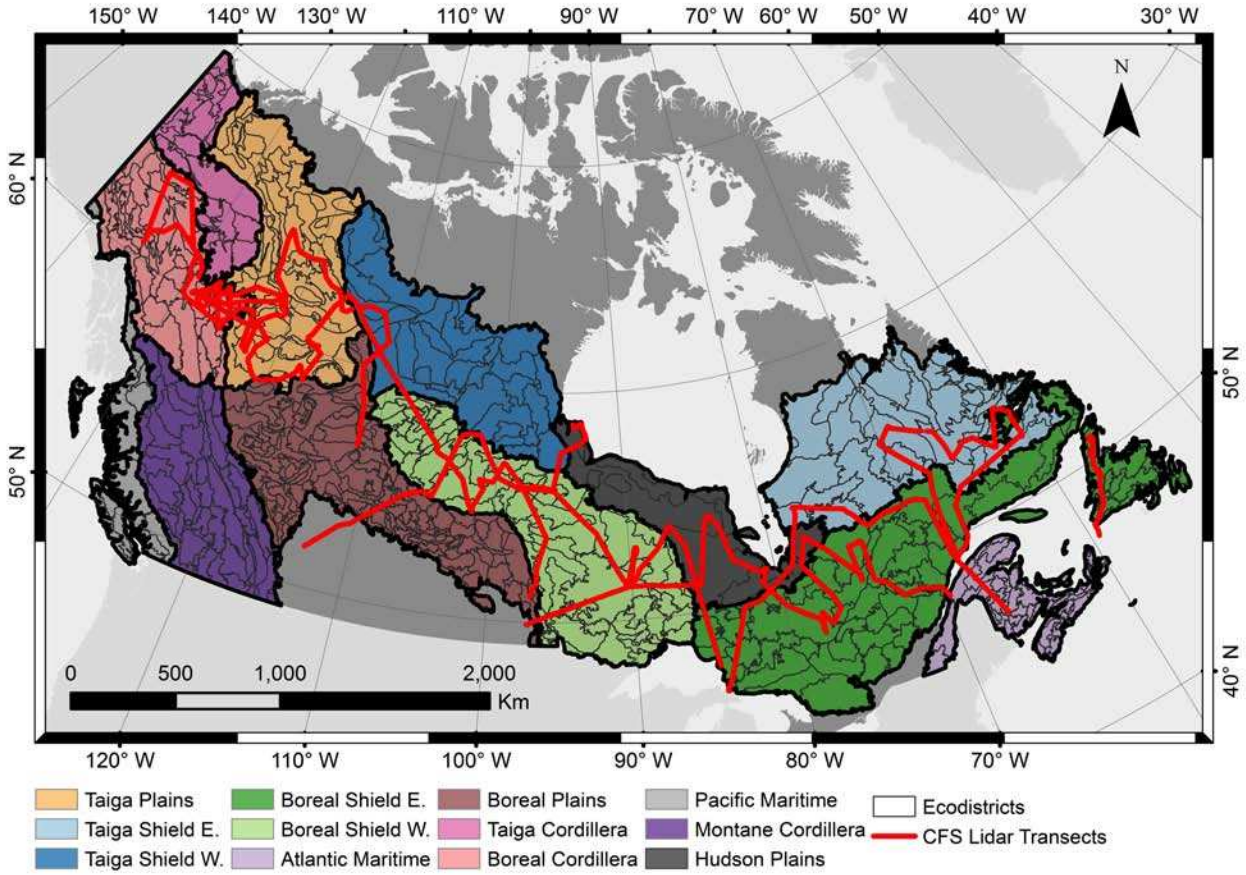


Figure 1: Map displaying the 10 forested ecozones of Canada and the ecodistrict subdivisions, along with flight path of the CFS Lidar transects

discrepancies at the ecozone level could signify a misrepresentation of certain forest conditions within the canopy height products. A linear regression was also performed between the height products at the ecodistrict level.

The height products were then compared against the 95th height percentile from the airborne Lidar data. To allow ecodistrict averages to correspond to the Lidar coverage, pixels that intersected the Lidar transects were identified for each height map and used to recalculate the ecodistrict averages. Linear regressions were then performed between the 95th height percentile and the height predictions at the ecodistrict level. Ecodistrict height difference was also calculated for each ecodistrict as:

$$Height\ Difference_i = Ht_{Map_i} - Lidar_{95th_i} \quad (1)$$

where Ht_{Map_i} is the height prediction and $Lidar_{95th_i}$ is the 95th height percentile for the i^{th} ecodistrict. The ecodistrict height differences were grouped into terrain roughness classes (i.e., standard deviation of elevation) to investigate the effect of slope on the accuracy of GLAS derived

height predictions. Boxplots were produced to display the height differences for each map product across the terrain roughness groups.

2.4 Investigating differences in biomass estimates

To explore the translation of differences in the GLAS derived height products to differences in carbon storage estimates, we calculated estimates of AGB for *Picea glauca* (white spruce), a common boreal species in Canada, from the ecodistrict height predictions. AGB can be estimated from tree height and diameter at breast height (DBH) using species-specific allometric equations. DBH was estimated for each ecodistrict using a Chapman-Richards model developed by Peng *et al.* (2001) that relates height and DBH for white spruce in Ontario boreal forests:

$$DBH_i = -\frac{\ln\left(1-\sqrt[c]{\left(\frac{H_i-1.3}{a}\right)}\right)}{b} \qquad \begin{matrix} a = 27.3476 \\ b = 0.0469 \\ c = 1.4360 \end{matrix} \qquad (2)$$

where DBH_i and H_i are the DBH and height estimates for the ith ecodistrict. Height values were again averaged to the ecodistrict level using only pixels that intersected the Lidar transects. The Lefsky (2010) height estimate for a Boreal Cordillera ecodistrict (32.9 m) was outside of the predictable range of the DBH/height model and was set to a height of 30.78 m and a DBH of 56.30 cm, the maximum white spruce dimensions used by Peng *et al.* (2001) in model development. The estimated DBH for an additional Lefsky (2010) ecodistrict in the Boreal Cordillera was outside of this range (65.30 cm) and was also set to 56.30 cm. Once DBH was derived for each ecodistrict, biomass components were calculated from the height and DBH values using allometric equations developed by Ung *et al.* (2008) for white spruce in Canada. Wood, bark, branch and foliage biomass components were summed for each ecodistrict to derive tree-level AGB estimates for white spruce. The same processing routine was applied to derive an AGB estimate from the 95th height percentile, and the following equation was applied to calculate the percent differences in AGB for each ecodistrict:

$$Percent\ Difference_i = \frac{AGB_{Map_i} - AGB_{\ Lidar_i}}{AGB_{Lidar_i}} * 100 \qquad (3)$$

where AGB_{Map_i} is the AGB estimate derived from each height map and $AGB_{\ Lidar_i}$ is the AGB estimate derived from the airborne Lidar data for the *i*th ecodistrict . The percent difference for each ecodistrict was grouped according to ecozone and displayed as a boxplot to investigate how height prediction differences could influence AGB estimates across ecozones.

3 Results

Figure 2 compares ecodistrict averages for the GLAS derived height products grouped according to ecozone. The height maps strongly agree in the Taiga Plains, Taiga Shield East, Boreal Shield East and West, Boreal Plains and Hudson Plains. Simard *et al.* (2011) had markedly higher height estimates in the Atlantic Maritime and Taiga Cordillera, while Lefsky (2010) had higher estimates in the Taiga Shield West, Boreal Cordillera, Pacific Maritime and Montane Cordillera. Height predictions for both maps were highest in the Pacific Maritime and Montane Cordillera and lowest in the Hudson Plains. Figure 3 displays the correlation at the ecodistrict level between a) the height products b) Lefsky (2010) and Lidar derived 95th height percentile and c) Simard *et al.* (2011) and Lidar derived 95th height percentile. The strongest correlation occurred between the height products

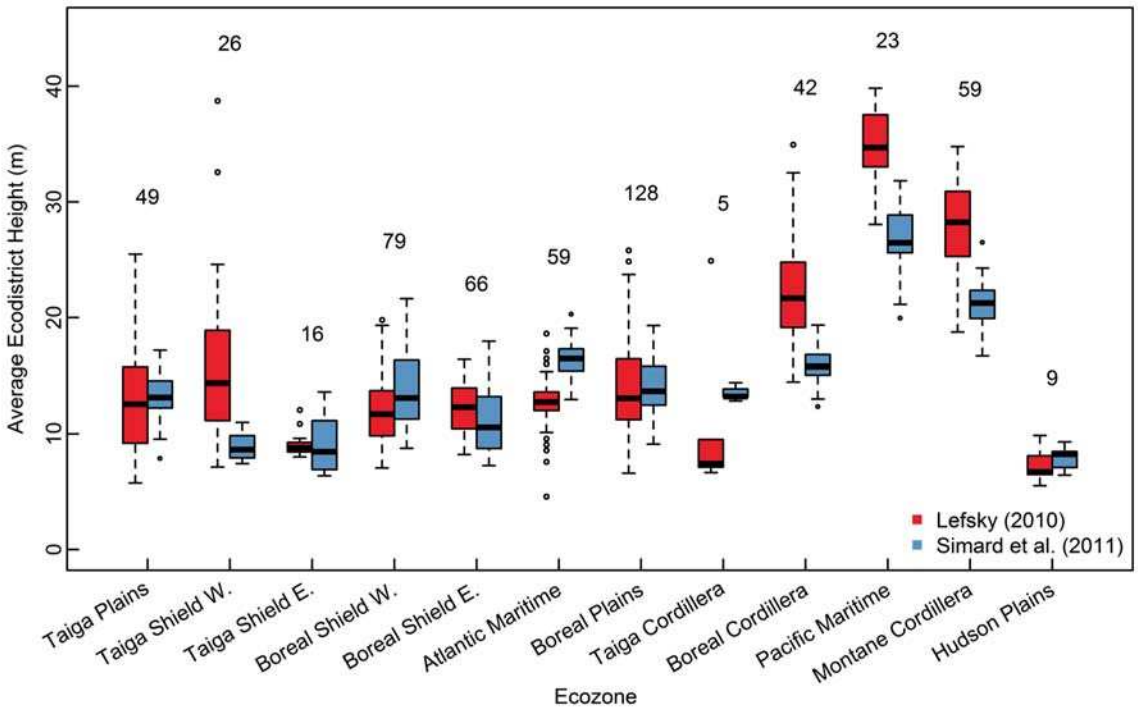


Figure 2: Ecodistrict average heights for Lefsky (2010) and Simard et al. (2011), grouped according to Ecozone. The number above each bin corresponds to the number of sampled ecodistricts with each ecozone.

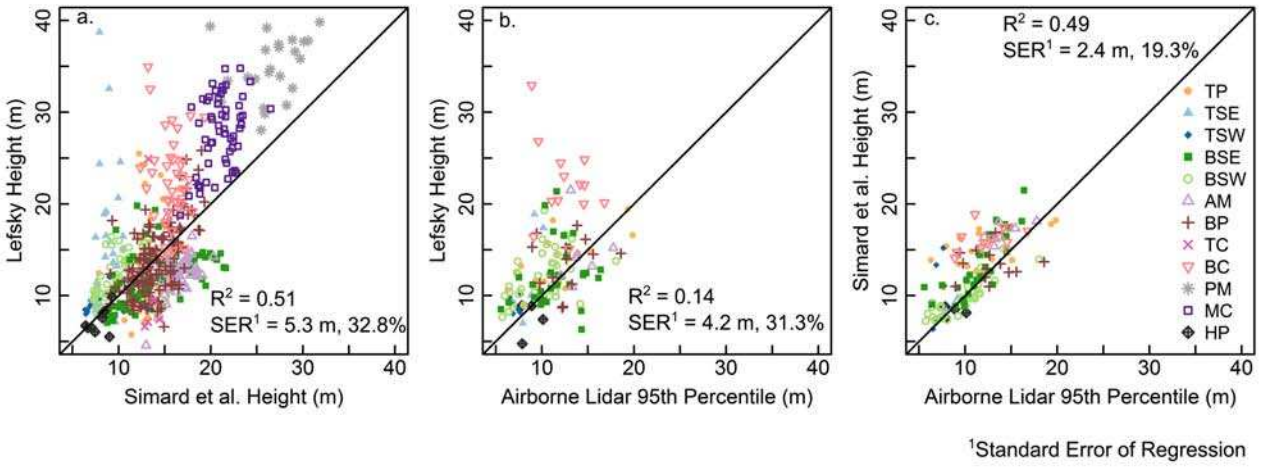


Figure 3: Relationships between a) Lefsky (2010) and Simard et al. (2011) b) Lefsky (2010) and the 95th height percentile from airborne Lidar c) Simard et al. (2011) and the 95th height percentile from airborne Lidar

(R^2 = 0.51), with the higher predictions for Lefsky (2010) apparent in the Boreal Cordillera, Montane Cordillera and the Pacific Maritime. Simard *et al.* (2011) had a stronger correlation to the 95th Lidar height percentile (R^2 = 0.49) than Lefsky (2010, R^2 = 0.14). The Lefsky (2010) map did predict taller canopies in most Boreal Cordillera ecodistricts when compared to the airborne Lidar data. Simard *et al.* (2011) also predicted taller canopies in most Boreal Cordillera ecodistricts, but to a lesser degree than Lefsky (2010). Simard *et al.* (2011) was less able to predict variation in height within the Taiga Plains than Lefsky (2010).

The height differences in relation to the 95th height percentile are displayed in Figure 4, with ecodistricts grouped according to terrain roughness (i.e., the standard deviation of elevation). The maps agree well with the airborne Lidar data when the standard deviation of elevation is low. When the standard deviation of elevation increases (i.e., more sloped terrain), both maps predict taller canopies than are predicted with airborne Lidar. The effect of terrain on height predictions appears to be greatest in the Lefsky (2010) map.

Figure 5 displays the percent difference for estimating white spruce AGB from each height product in comparison to estimates derived from the 95th height percentile. Only ecozones that contain more than 10 sampled ecodistricts were included in this analysis. AGB estimates had the strongest agreement in the Boreal Shield West and the Boreal Plains, while AGB estimates were higher in the Boreal Cordillera, most significantly by Lefsky (2010). For Lefsky (2010), the two ecodistricts in the Boreal Cordillera that were outside of the predictable range for the height/DBH relationships had > 30% difference for predicting AGB. Lefsky (2010) also predicted higher AGB in the Boreal Shield East while Simard *et al.* (2011) predicted higher AGB in the Taiga Plains. With the exception of the Lefsky (2010) AGB estimates for the Boreal Cordillera, the mean percent difference for AGB predictions was < 2% for all ecozones.

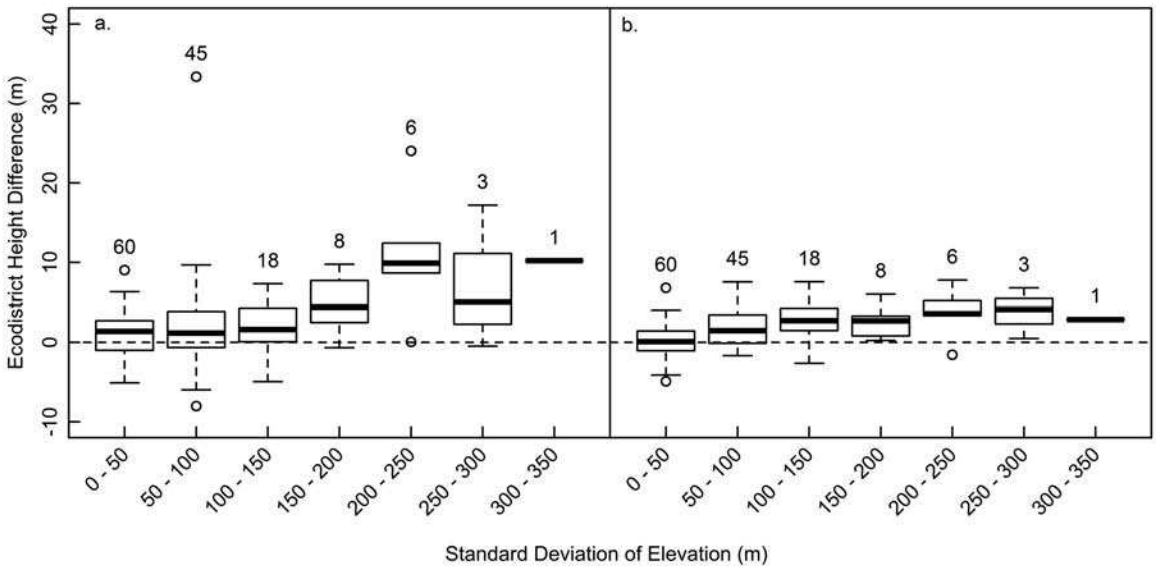


Figure 4: Ecodistrict height difference for a) Lefsky (2010) and b) Simard et al. (2011). Ecodistricts are grouped by the standard deviation of elevation within each ecodistrict

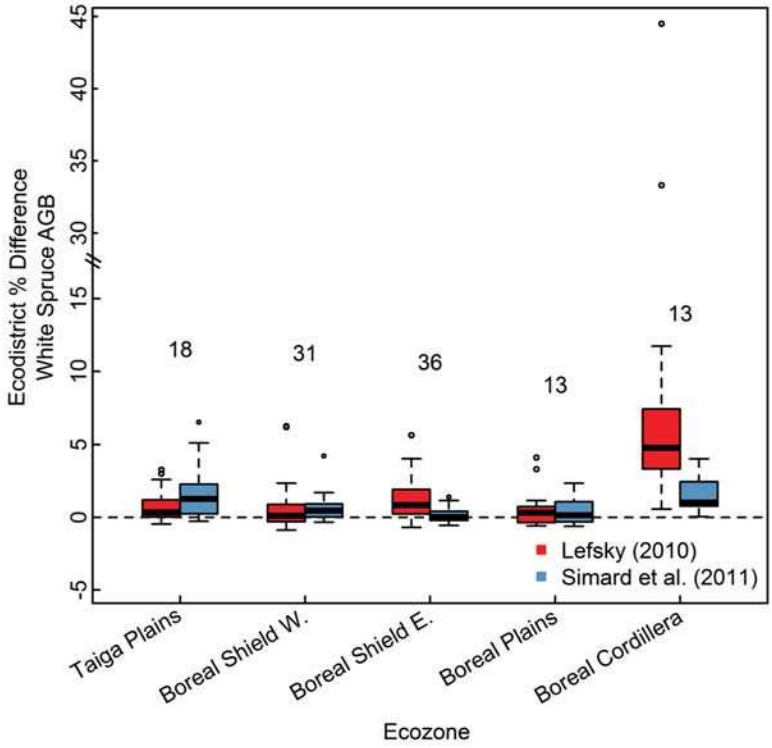


Figure 5: Ecodistrict percent difference for predicting AGB with height estimates for Lefsky (2010) and Simard et al. (2011), grouped by ecozone

4 Discussion

Comparing GLAS derived canopy height products to height measurements from airborne Lidar data allows the assessment of the usefulness of GLAS data for predicting canopy height at the ecozone and ecodistrict level in Canada. While globally distributed, GLAS does not provide spatially exhaustive coverage, as such accurate methods of extrapolation are necessary to produce wall-to-wall global height maps from GLAS data. Large discrepancies between the Lefsky (2010) and Simard *et al.* (2011) height maps in the Pacific Maritime, Montane Cordillera and Boreal Cordillera highlight the influence that processing and extrapolation techniques can have on GLAS derived height predictions and carbon storage estimates. The stronger relationship between Simard *et al.* (2011) and Lidar derived heights as compared to Lefsky (2010) suggests that climate and ancillary indicators have some added value for extrapolating GLAS data across Canada. Alternatively, MODIS spectral data does appear more sensitive to variations in height in the Taiga Plains than the indicators used by Simard *et al.* (2011), as the Simard *et al.* (2011) height predictions do not capture the height variability observed with airborne Lidar. As differences in the methodologies used by Lefsky (2010) and Simard *et al.* (2011) extend beyond the chosen indicators, differences in accuracy should not be entirely attributed to the suitability of these indicators. The segment based

approach applied by Lefsky will also result in an amplification of particular values, as the same height will be applied to a particular segment, with greater opportunity for variation in pixel-based products, such as from Simard.

Research shows that the ability to estimate forest height with GLAS decreases as terrain slope within the laser footprint increases (Duncanson *et al.* 2010; Xing *et al.* 2010; Pang *et al.* 2011). As the slope within a laser footprint increases, the vertical range of the Lidar waveform increases, which can cause an overestimation of canopy height from the waveform (Lefsky *et al.* 2005). This provides an explanation for why both height maps predicted taller canopies when the standard deviation of elevation is large within an ecodistrict. The larger effect of slope on Lefsky (2010) height predictions is likely the reason for the discrepancies between the Lefsky (2010) and Simard et al. (2011) in the mountainous regions of the Pacific Maritime, the Boreal Cordillera and the Montane Cordillera.

The translation of error from GLAS derived height products to estimates of AGB is of particular interest for carbon modeling activities. Carbon storage in forest AGB is a product of three fundamental forest attributes: tree height, DBH, and stem density. The Lefsky (2010) and Simard *et al.* (2011) height maps predict only one of these fundamental attributes, but represent a potential large source of error for estimating carbon budgets. We have shown that the height differences in the GLAS derived products translate to small differences in AGB at the ecozone scale, with the exception of the Boreal Cordillera. The taller canopies that were predicted using the Lefsky (2010) product in the Boreal Cordillera in relation to the airborne Lidar data could lead to an AGB difference of ~5% across the ecozone.

5 Conclusion

The GLAS derived height products were able to characterize forest height within most Canadian ecozones while they had variable success at the finer scale ecodistrict level. The Simard *et al.* (2011) height product had a stronger correlation to the 95th height percentile from airborne Lidar at the ecodistrict level, while Lefsky (2010) predicted taller canopies than airborne Lidar in most Boreal Cordillera ecodistricts. From the stand point of carbon modeling, we must determine if accurate estimates of height and AGB at the ecozone scale are sufficient for modeling activities or if finer scale variations are required. For instance, if management activities vary spatially within an ecozone (i.e. more management in the south portion of the ecozone), understanding the within ecozone variations in height and AGB may be necessary to accurately model forest C dynamics. Large discrepancies between the two GLAS derived height products in the Boreal Cordillera, Pacific Maritime and Montane Cordillera reveal the impact that processing and extrapolation techniques have on the final height estimates, which would translate to potentially large differences in AGB estimates for carbon models. Despite differences in processing techniques, both GLAS derived height predictions predict taller canopies than airborne Lidar over sloped terrain. While these current analyses are constrained to the ecozone level, we plan future work at spatial resolution of each global height product, through custom generalization of the lidar data.

Acknowledgements

We would like to acknowledge Canadian Forest Service funding of the national lidar trasects. The authors of the global forest height products are thanked for the efforts in producing these important products and for sharing the results in a free and accessible fashion.

References

Boisvenue, C. & Running, S.W. (2006). Impacts of climate change on natural forest productivity - evidence since the middle of the 20th century. *Global Change Biology*, doi:10.1111/j.1365-2486.2006.01134.x

Denman, K.L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P.M., Dickinson, R.E., *et al.* (2007). Couplings between changes in the climate system and biogeochemistry, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.

Dubayah, R.O. & Drake, J.B. (2000). Lidar remote sensing for forestry. *Journal of Forestry*, *98*, 44–46.

Duncanson, L.I., Niemann, K.O., & Wulder, M.A. (2010). Estimating forest canopy height and terrain relief from GLAS waveform metrics. *Remote Sensing of Environment*, *114*, 138–154.

Flannigan, M.D., Logan, K.A., Amiro, B.D., Skinner, W.R., & Stocks, B.J. (2005). Future area burned in canada. *Climatic Change*, *72*, 1–16.

Gillis, M.D., Omule, A.Y., & Brierley, T. (2005). Monitoring Canada's forests: the National Forest Inventory. *The Forestry Chronicle*, *81*(2), 214–221.

Goetz, S. & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Management*, *2*(3), 231–244.

Houghton, R.A., Hall, F., & Goetz, S.J. (2009). Importance of biomass in the global carbon cycle. *Journal of Geophysical Research*, doi:10.1029/2009JG000935

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195–213.

Kane, V.R., McGaughey, R.J., Bakker, J.D., Gersonde, R.F., Lutz, J.A., & Franklin, J.F. (2010). Comparisons between field- and LiDAR-based measures of stand structural complexity. *Canadian Journal of Forest Research*, *40*, 761–773.

Le Quéré, C., Raupach, M.R., Canadell, J.G., Marland, G., Bopp, L., Ciais, P., *et al.* (2009). Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, doi:10.1038/ngeo689

Lefsky, M. a. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, *37*(L15401), 1–5.

Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, *70*, 339–361.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Del Bom Espirito-Santo, F., *et al.* (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, *32*.

Magnussen, S. & Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, *28*, 1016–1031.

Masek, J.G., Cohen, W.B., Leckie, D., Wulder, M.A., Vargas, R., de Jong, B., *et al.* (2011). Recent rates of forest harvest and conversion in North America. *Journal of Geophysical Research*, doi:10.1029/2010JG001471

McGaughey, R. J. (2012). FUSION/LDV: Software for LIDAR data analysis and visualization. Pacific Northwest Research Station, Forest Service, United States Department of Agriculture. http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf. Accessed 15 May 2012.

Natural Resources Canada (2011). *The state of Canada's forests. Annual report 2011*. Ottawa: Natural Resources Canada, Canadian Forest Service.

Pang, Y., Lefsky, M., Sun, G., & Ranson, J. (2011). Impact of footprint diameter and off-nadir pointing on the precision of canopy height estimates from spaceborne lidar. *Remote Sensing of Environment*, *115*, 2798–2809.

Peng, C., Zhang, L., & Liu, J. (2001). Developing and validating nonlinear height – diameter models for major tree species of Ontario ' s boreal forests. *Northern Jounral of Applied Forestry*, *18*(3), 87–94.

Rabus, B., Eineder, M., Roth, A., & Bamler, R. (2003). The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*, 241–262.

Simard, M., Pinto, N., Fisher, J.B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research*, *116*(G04021), 1–12.

Ung, C., Bernier, P., & Guo, X. (2008). Canadian national biomass equations: new parameter estimates that include British Columbia data. *Canadian Journal of Forest Research*, *38*, 1123–1132.

Wulder, M.A., Campbell, C., White, J.C., Flannigan, M., & Campbell, I.D. (2007). National circumstances in the international circumboreal community. *Forestry Chronicle*, *83*(4), 539–556.

Wulder, M.A., White, J.C., Cranny, M.M., Hall, R.J., Luther, J.E., Beaudoin, A., *et al.* (2008). Monitoring Canada's forests. Part 1: Completion of the EOSD land cover project. *Canadian Journal of Remote Sensing*, *34*(6), 549–562.

Wulder, M.A., J.C. White, C.W. Bater, N.C. Coops, C. Hopkinson, and G. Chen. (*In press*). Lidar plots—a new large-area data collection option: context, concepts, and case study. Canadian Journal of Remote Sensing. (Accepted Jun 14, 2012.)

Xing, Y., de Gier, A., Zhang, J., & Wang, L. (2010). An improved method for estimating forest canopy height using ICESat-GLAS full waveform data over sloping terrain: A case study in Changbai mountains, China. *International Journal of Applied Earth Observation and Geoinformation*, *12*, 385–392.

Zwally, H.J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., *et al.* (2002). ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics*, *34*, 405–445.

Demonstrations of a Successful LiDAR Business Case – Conception to Reality

Tony Brown[1]
Mike Sutton[2]

[1]Forests NSW, Australia tony.brown@sf.nsw.gov.au
[2]Forests NSW, Australia mike.sutton@sf.nsw.gov.au

Paper Number: SL2012 - 064

Abstract

In a challenging business environment with pressures to maintain service with less staff, improving safety and lowering costs, Forests NSW (FNSW) faced a challenge to produce a financially sound business case to capture aerial LiDAR over large areas of commercial native forest. Early research work, two externally funded LiDAR campaigns and large volumes of free data supplied by local government provided the catalyst for building the business case.

FNSW undertakes timber harvesting under a complex and costly regulatory regime, within a highly variable native forest estate, under high levels of external scrutiny. To construct a financially attractive business case for LiDAR acquisition, FNSW needed to focus on areas where significant financial savings could be quickly realised – the obvious candidate being reduction in the amount of field work.

In large part, current regulation requires operational management of watercourses as mapped on published topographic maps, rather than as found on ground. Traditionally FNSW mark all stream buffers with paint, a slow and expensive process. Using LiDAR derived streams and applying current regulatory constraints to GIS features, streamside reserves with buffer widths greater than 20m are now 'marked' by hand held GPS located in harvesting machines. Through use of LiDAR and in-harvester GPS, significant financial savings are consistently demonstrated on every harvesting operation. In addition to efficiency gains from stream network mapping, improvements in quantifying volume and identifying the location of commercial quantities of resource have also been achieved by using LiDAR-derived products. Using basic DTM products such as streams, slopes and shaded relief, as well as DSM data including canopy height models and a locally derived vegetation stratification, significant savings have been demonstrated in operational planning. Improved product volume estimates, better contractor scheduling, quicker planning times, improved road planning through better resource and terrain knowledge, improved safety outcomes and informed, targeted field work all contribute to significantly more informed and improved harvest plans and operational management.

In building an LiDAR acquisition business case FNSW focused on areas of high cost which could be significantly reduced through application of LiDAR. Acquisition costs have been recouped well ahead of forecast schedule, further improving an originally attractive NPV. Field work is expensive and often unsafe; however using simple derived products like streams, slope and shaded relief, costs can be significantly reduced while also improving safety performance. Further refinement of research results is not necessarily the only way to demonstrate the benefits of LiDAR. A sound business case, backed up by in-field testing, operator feedback and local adaptation to local conditions can provide evidence needed to justify what are still high level costs required to implement LiDAR at the operational scale. The case for LiDAR can be further enhanced by using it in as many parts of the organisation as possible, demonstrating its effectiveness and continually improving the return on the initial capital investment though operations-driven product development. Future challenges to the operational use of LiDAR include identifying cost-effective approaches for keeping datasets current.

1. Introduction

In a challenging business environment with pressures to maintain service levels using less staff, while at the same time improving safety and lowering costs, Forests New South Wales (FNSW) faced a challenge to produce a financially sound business case to capture LiDAR over large areas of commercial native forest. Despite promising results from significant research work identifying the potential benefits of using LiDAR data for operational forest management, senior management were reluctant to approve spending on LiDAR data without a financially sound, detailed business case. To date, there has been little published material on preparation of such business cases for large scale economically justified LiDAR campaigns for operational forest management.

The first application of LiDAR at an operational scale by FNSW occurred following a large investigation into forest management and conservation tenure in FNSW Western Region. As part of this process, external funding was made available to identify areas suitable for young age silviculture in White Cypress Pine (*Callitris glaucophylla*). Through an internal process LiDAR emerged as the tool of choice for this process, with a large capture campaign initiated in 2007 for 242,000 ha of native forest estate. Following successful application of this technology as an operational tool in Western Region, coastal regions of FNSW began to seek access to this technology. LiDAR data was sourced from existing local datasets, with LiDAR data provided by local government agencies further contributing to operational testing of large scale operational LiDAR data in FNSW Central region.

FNSW undertakes timber harvesting under a complex and costly regulatory regime, within a highly variable native forest estate, under high levels of external scrutiny. To construct a financially attractive business case for LiDAR acquisition, FNSW needed to focus on areas where significant financial savings could be quickly realised – with the obvious candidate being reduction in the amount of field work. Opportunities to reduce field work through operational efficiencies were to be investigated, with both ground and non-ground data targeted to identify areas of significant savings from which FNSW would be able to construct a financially attractive and operationally sound business case for internally funded LiDAR capture over its coastal native forest estate.

2. Method

The challenge facing FNSW was to demonstrate an improvement in the efficiency of operational practices which would ultimately result in reduced management costs, and at the same time deliver significant improvements in employee safety. Given the scale of the project, a replicated study comparing traditional planning methods with LiDAR derived data across several operational management areas was not feasible. Rather, a business case was prepared using estimates of potential planning and supervision savings and efficiencies made by experienced forestry staff.

At the time of business case development, FNSW was undergoing substantial restructuring within the native forest workforce. Initial estimates around time savings likely to be realised after LiDAR implementation were reflected in significantly lower staff numbers in key positions after the restructure, and are detailed in Table 1. There was no corresponding reduction in the planning or supervision workload. In fact as a result of market conditions, lower than average yields/ha and higher than usual occurrences of wet weather, the number of plans required and the number of contractors to supervise increased slightly over the subsequent 12 months. Therefore, the ultimate demonstration of the success of LiDAR implementation was being able

to maintain business operations with reduced staff resources following business restructure.

Table 1: Pre and Post Restructure staffing levels.

Team	Pre Restructure Staffing Levels	Post Restructure Staffing Levels
Operational Planning	10	7
Operational Supervision	22	17

Data for the Central Region study were provided to FNSW as .xyz files classified as ground and non-ground returns. The ground data had been 'thinned'; however no indication of the extent or processing had been given. Further reductions in data strike intensity had been requested by the local government agencies sponsoring LiDAR capture over non-urban areas, which typically include the FNSW estate. Notwithstanding these limitations and cognisant of the impacts they may have, FNSW processed the data into 1m Ground Sample Distance rasters for Digital Terrain Model (DTM) and 1m Digital Surface Model (DSM) using ERDAS Imagine® Terrain Preparation software. Typical ground returns varied significantly across the datasets, averaging around 1.5 returns per square metre.

Preliminary investigations into LiDAR data to augment operational planning and harvest supervision processes were undertaken, with initial products detailed in Table 2. All derived datasets and spatial products were produced using ESRI ArcGIS® tools. Further investigations into field implementation of Global Positioning System (GPS) units within harvesting machinery to manage operational exclusion zones informed by LiDAR derived datasets would also be conducted.

Table 2: Initial LiDAR derived data provided to operational staff, including delivery format.

Product	GSD (m)	Description	Delivery Format
Shaded Relief	1	Morning shaded relief of ground surface	Raster dataset
Slope	1	Degree slope as classified integer to meet operational slope classes	Raster dataset
CHM	1	DSM – DEM	Raster dataset
Strata	5	Integration of 2 inputs analysed over 45 x 45 m window. 1. Classified mean CHM 2. A count of pixels > 30 m Output classified into 5 'productivity' classes	Raster dataset
Streams	1	ArcGIS derived D8 surface flow accumulation threshold at 0.5 ha and converted to polylines	Polyline

2.1 Study area

The study area was a subset of FNSW Central Region, an estate area of just over 50,000 ha of the Mid North Coast of New South Wales (see Figure 1.) Multiple datasets were used within the study area, all of which were captured by the Hastings-Port Macquarie and Taree Local Government agencies in 2006 through external contract, and provided to FNSW under a minimal restriction data use licence agreement.

3. Results

Despite the low point intensity (< 2/m^2) over the trial site which resulted from lower capture intensities and point thinning, FNSW staff were very impressed by the quality of output data. Shaded relief imagery alone provided planning staff with significantly improved information about planning areas, allowing staff to focus all ground work on areas of known issues (such as

stream crossings, road and log dump locations), rather than 'exploring' an area on foot or using aerial photography to find issues. Both morning and evening shaded relief imagery were produced, providing different perspectives of the ground surface to highlight all features.

Slope was initially calculated on a 1m Ground Sample Distance (GSD) surface and found to be very useful, although perhaps too detailed, for operational planning. Subsequent discussion with operational staff, analysis of slope products and evaluation of alternative processing options, resulted in a change to a 5 m GSD raster with a 3x3 cell focal mean. This provided a smoothed output which was significantly more representative of the area from an operational forest management perspective.

Streams were derived using ArcGIS surface tools, which implement a D8 surface flow analysis (Jenson and Domingue 1988). All streams were thresholded at 0.5 ha, and converted to lines which were consolidated into a single ESRI Geodatabase which is shared to all GIS users. As existing regulation of harvesting adjacent to streams requires management 'as mapped' on published topographic maps, a "stream order" field was included within the LiDAR derived stream vectors to enable operational staff to attribute the LiDAR derived stream with the order as defined on existing topographic map sheets.

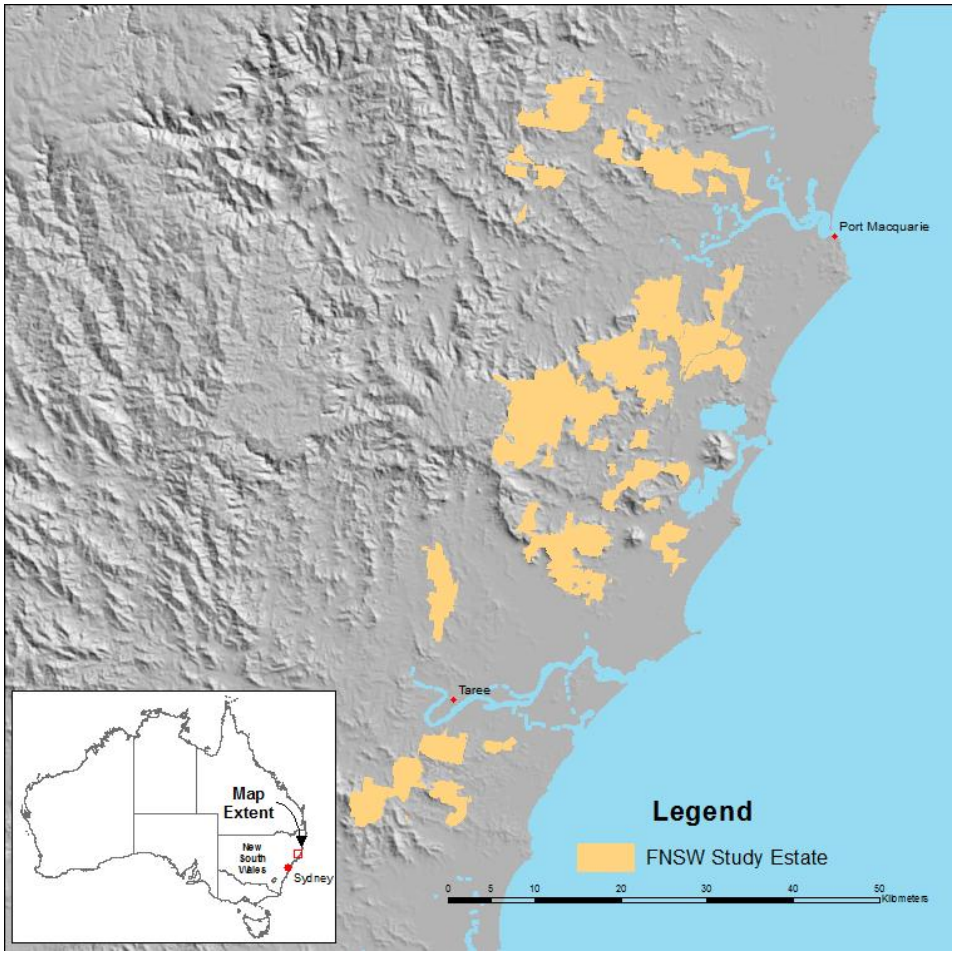


Figure 1: Study area location.

This manual editing process, which while time consuming for operational planning staff, provides an accurate stream location while including regulatory requirements for stream order. This information can then be uploaded to a standard GPS unit installed in harvesting machinery. Inclusion of a stream bed width value ensures that centreline outputs from the D8 process meet

regulatory requirements by including stream width. Streams with a non-harvest zone >20 m are now marked electronically using LiDAR derived stream locations and in-harvester GPS, greatly reducing the manual field marking requirements within harvest compartments.

While FNSW found a simple D8 surface flow analysis identified nearly all on-ground watercourse features requiring management during forest operations, there are errors typically associated with road/watercourse interfaces. When mapping larger minimum catchment sizes (eg. >10 ha), these errors can be managed by re-sampling to larger DTM cell sizes (Mouton 2005). However for operational management, smaller catchments are required; and with no automated method of 'conditioning' DTM for these errors, manual editing of the point cloud or DTM is an expensive solution. As such, FNSW currently uses manual correction of D8 derived modelled streams.

Simple training was conducted for FNSW field staff and contract harvest staff to explain how LiDAR technology works, how to interpret the outputs and how to use the information with GPS units. A range of in-car, hand-held and agricultural GPS were trialled for ease of use and data management capability. The Garmin 62 series hand held proved the most widely accepted for all applications, limiting image input requirements to a single data type. A rapid acceptance and uptake of the combined technology by both FNSW field staff and contractors, coupled with initial focused supervision from FNSW staff to ensure compliance and manage contractor queries, has significantly reduced in field boundary marking requirement for FNSW.

In addition to efficiency gains from stream network mapping, improvements in quantifying volume and identifying the location of commercial quantities of resource have LiDARo been achieved through the use of LiDAR-derived products. Initially, CHM data was made available for interpretation as continuous surface. On-screen interpretation proved highly valuable for operational planning, but had limitations when used for pre-harvest resource assessment. Based on feedback from operational staff, further analysis of CHM data was undertaken using a range of methods including focal mean, focal sum, measures of variation and combinations of these processes. The final output involved two reclassifications of the data within a 45 m x 45 m focal window: a mean canopy height and a count of pixels over a 30 m threshold. These outputs were combined into a single dataset and classified into 5 structural strata, and which were known as Native Forest Strata. This output proved a significant improvement over existing aerial photographic interpretation forest typing, which was variable in nature and often out of date as a result of forest harvesting, fire or stand growth.

The Native Forest Strata was widely used by operational staff for nearly 12 months and worked very well for a range of complex native forest types, however refinement was required in lower productivity strata. Further investigation resulted in the use of a 45 m x 45 m focal window Sum of Squares of the 1m GSD CHM data, which was classified into 8 Forest Productivity strata. This output has replaced the original Native Forest Strata, and is proving more informative across the full range of forest types on the FNSW coastal estate.

Initially, all data was provided as stand-alone raster products, and loaded into ArcGIS raster catalogs. These datasets were then served out over existing local area networks via layers to all operational staff. Subsequent developments within ESRI ArcGIS led FNSW to change to mosaic datasets. The advantage of this approach is that much of the processing of derived data such as shaded relief can be done 'on the fly' using integrated mosaic dataset functions. This reduces the amount of pre-processing required, simplifies data updates and allows for better temporal data management. Table 3 details the revised suite of LiDAR derived products and their delivery format.

Table 3: Revised LiDAR derived data provided to operational staff after feedback and product refinement, including delivery format.

Product	GSD (m)	Description	Delivery Format
Shaded Relief	1	Morning and afternoon shaded relief of ground surface	DEM referenced Mosaic dataset
Slope	5	5 m GSD DEM derived slope raster with 3 x 3 cell focal mean	Mosaic dataset
CHM	1	Combined shaded relief CHM and semi transparent continuous CHM as grouped layers	Mosaic dataset
Strata	5	45 m x 45 m focal sum of 1 m GSD CHM^2, re-sampled to 5 m GSD and classified into 8 classes.	Mosaic dataset
Streams	1	ArcGIS derived D8 surface flow accumulation threshold at 0.5ha and converted to polylines	Polyline

4. Discussion and Conclusions

It became apparent soon after implementation of LiDAR data and the field use of derived products within Central Region just how informative and time-saving these can be during forest planning and operational management. Even the low strike intensity data provided significant opportunity for interpretation by operational planning and field staff. Rapid uptake of the technology by office and field staff highlights opportunities where individuals identified ways to improve their knowledge or efficiency of current tasks.

There have been significant improvements in the production and quality of operational plans. This is evidenced by a reduction in the number of plan amendments required after plan certification, reduced time for plan preparation and the improved estimation of harvest duration times due to more accurate volume assessments. Better knowledge of the spatial location of the resource across compartments has resulted in more appropriate road, log dump and snig track locations, improving contractor harvesting efficiencies and reducing the costs of road maintenance and upgrading for harvesting.

Integration of stream data and GPS in harvesting machines was the cornerstone of the business case for LiDAR acquisition. It was clearly evident through integration of these technologies that significant staff savings could be realised by increasing the number of operations individual supervisors had responsibility for. Despite not replacing staff lost through natural attrition, FNSW is now operating with increased confidence in a complex regulatory environment with a significantly more streamlined workforce.

While not the only contributor, implementation of freely acquired LiDAR data into business operations is a major component of FNSW success in maintaining business standards with significantly fewer staff. A business case for acquisition of further LiDAR data was formulated around savings derived through efficiencies as well as significant cost avoidance, which was subsequently approved with data capture now complete. While much of the financial data supporting the business case are commercial in confidence, a summary is presented in Table 4, which shows major components of the original business case as well as updated figures developed during a project review derived from actual costs and savings.

The application of these new approaches has generated significant discussion with industry and with regulators. FNSW has begun working with regulators and LiDAR researchers to investigate methods to provide a regulatory environment which integrates LiDAR and GPS technology, with the objective of providing improved environmental outcomes within a more efficient regulatory and compliance framework. In this context, conditioning of DTM around

road/stream interfaces becomes critical. Identification of non-harvest zones around streams is likely to require catchment area as one of the input variables. This will require accurate mapping of entire catchments to maintain regulatory uniformity and provide confidence in any changes to the forest regulatory system.

Table 4: Financial summary of major components for LiDAR acquisition business case and revised business case financials at conclusion of project in AU$.

	Description	Yr 1	Yr 2	Yr 3	Yr 4	Yr 5
Business Case	Costs including Data capture, processing and management, project management and GPS hardware	$550,000	$189,000	$15,000	$10,000	$5,000
	Benefits including more efficient harvest planning, boundary marking	$175,000	$350,000	$450,000	$450,000	$450,000
	Annual NPV	-$341,000	-$208,000	$119,000	$419,000	$696,000
	Total NPV	**$696,000**				
Business Case Review	Costs including Data capture, processing and management, project management and GPS hardware	$403,000	$390,000	$10,000	$5,000	$5,000
	Benefits including more efficient harvest planning, boundary marking	$200,000	$450,000	$600,000	$700,000	$700,000
	Annual NPV	$582,000	$612,000	$888,000	$1,182,000	$1,450,000
	Total NPV	$1,450,000				

The success of the operational application of LiDAR is dependent on seeking, and acting on, feedback from operational staff on the use of various derived products. While simple training in the technology, its use and limitations was undertaken for all users, regular feedback is also sought to identify where improvements can be made from an operational perspective. For example moving from the 1 m GSD slope data to a 5 m GSD with a 3 x 3 cell focal mean has significantly improved the operational use of slope data, informing discussions with harvesting contractors on slope classes within current compartments, and providing quantitative information for tendering on future operations.

Additionally, keeping abreast of changes in software allows significant improvements in turnaround from contractor data delivery to internal 'publication' of derived products for operational staff. An example of this is improvements from ArcGIS 9.3 to ArcGIS 10.0 and the adoption of mosaic datasets. These changes have reduced data processing times, improved temporal data management and increased the number and range of data products available to operational staff.

Rather than wait for development of a statistical relationship between LiDAR and plot data, FNSW has opted for incremental improvements in the use of native forest stratification, to provide significant efficiency improvements and cost savings. The first stratification method provided staff with gains over simple CHM interpretation, and when further improvement was needed, analysis was undertaken, field tested and a simple product developed and provided, all with a short turnaround to meet the operational needs of field staff. During development of the secondary stratification method, FNSW modified its strategic and tactical inventory plot location specifications to enable plots to be linked to LiDAR data. FNSW is currently investigating correlations between plot data and LiDAR, hoping to further improve volume and product estimation using LiDAR. Future work will include review of the Native Forest

Inventory Framework to incorporate LiDAR, and considerations of cost-effective alternatives to keep CHM datasets current in the dynamic forest estate.

Through investigation of freely available data sets, albeit of a lower than deemed optimal strike intensity, FNSW was able to build a financially attractive LiDAR acquisition business case. FNSW has subsequently increased its LiDAR coverage, through both internal acquisition and through access to whole of government data sets, to now have well over 1,000,000 ha of productive forest estate under LiDAR, and planned acquisition of a further 500,000 ha.

Acknowledgements

Many thanks to the staff of FNSW Native Forest Directorate, especially those from Western and Central Regions for their patience, understanding and willingness to use of LiDAR data.

References

Jenson, S. K. and Domingue, J. O. 1988. Extracting topographic structure from digital elevation data for geographic information systems analysis. Photogrammetric Engineering and Remote Sensing. Vol. 54, No. 11, pp. 1593-1600

Mouton, A. Generating Stream Maps Using LiDAR Derived Digital Elevation Models and 10-m USGS DEM. M.Sc. thesis University of Washington, 2005.

The influence of scan mode and circle fit algorithms on the extraction of stem diameter and volume from TLS data

P. Pueschel[1], G. J. Newnham[2], G. Rock[1], T. Udelhoven[1], W. Werner[3], J. Hill[1]

[1]University of Trier, Department of Environmental Remote Sensing and Geoinformatics, Trier, Germany, (puesch, rock, udelhoven, hillj)@uni-trier.de
[2] CSIRO Land and Water, Clayton South, Australia, glenn.newnham@csiro.au
[3]University of Trier, Department of Geobotany, Trier, Germany, werner@uni-trier.de

Paper Number: SL2012-068

Abstract

Terrestrial laser scanning (TLS) has been used to estimate a number of structural forest parameters, including LAI, crown cover, tree height, and diameter at breast height (DBH). A number of studies have shown that DBH, which is a main parameter in forest inventory, can be successfully extracted from TLS data. However, due to differences in the experimental set-up, data processing and test plot characteristics, the reported estimation accuracies varied strongly. In order to provide consistency and maximize accuracy, a systematic study of the influence of these variables is required. To contribute to such an approach, we focused our study on the effects of the scan mode (single vs. multiple scans) and circle fit algorithms on automated stem detection, stem diameter and volume extraction. Results show that tree stems which are unobstructed can be detected with mean accuracies of 94 % to 96 %. The extracted DBH agree closely with reference tape measurements, with single scan RMSE's ranging from 1.39 cm to 2.43 cm. With regard to the different circle fit algorithms tested (Lemen, Pratt, Taubin), the Lemen algorithm showed the best overall performance. The DBH estimation based on merged scan data (multiple scans) performed superior compared to the single scan data, with RMSE's ranging from 0.66 cm to 1.21 cm. The differences between single and merged scan mode are reflected in the stem volume estimates, too. Single scan stem volumes exhibit a large variability with mean deviations from the reference volumes ranging from -18% to 25%. By contrast volumes extracted from the merged scans only vary weakly (-1% to 6%).

1. Introduction

The application of terrestrial laser scanning (TLS) for forest inventory is still a relatively new field of research. This is due to high scanner costs and limited capacities for an automated data processing. However, encouraged by the promising results of pilot studies (e.g. Simonse et al., 2003, Thies & Spiecker, 2004) and boosted by the rapid advance in scanner technology, forestry-related research and applications of TLS have been on the increase in recent years. To date, TLS studies have covered a number of biophysical and structural forest parameters including tree location (e.g. Simonse et al., 2003), tree height (e.g. Maas et al., 2008), diameter at breast height (e.g. Bienert et al., 2006), stem volume (e.g. Tansey et al., 2009), biomass (Holopainen et al., 2011, Yao et al., 2011), Leaf Area Index and gap fraction (e.g. Jupp et al., 2008). While these studies confirmed the basic potential of TLS to successfully estimate these parameters, a systematic study of the factors which influence the information extraction from TLS data is still lacking. Possible influences are numerous and include:

- Ranging method (e.g. time-of-flight discrete return, continuous wave phase-shift)
- Scanner characteristics (e.g. laser wavelength, beam divergence)
- Scan settings (e.g. angular resolution, integration time)

- Scan mode (e.g. single or multiple scans)
- Data processing (e.g. filtering, higher level algorithms for deriving structure)

These influences exhibit interdependencies which for forest applications are further complicated by different stand and plot characteristics (e.g. canopy layering, degree of undergrowth, stem density, slope). As a consequence, the existing studies yield widely different results, even for such seemingly simple tasks as stem detection and diameter extraction. The basic differences between these studies also make it difficult to objectively compare the proposed approaches, a fact which calls for a more systematic study into the TLS-related influences. The present study intends to contribute to such an approach by assessing two of these influences, namely circle fitting and scan mode, on the stem detection, stem diameter and volume extraction from phase-shift TLS data. Other important influences will be addressed in a follow-up study.

2. Methods

2.1 Study area

This study was carried out as part of a larger campaign with the purpose of validating forest structural metrics derived from airborne hyperspectral and laser scanning data, and which serves as preparation for the German hyperspectral satellite mission EnMAP (http://www.enmap.org/). The study area is the *Pfälzerwald* forest near Kaiserslautern, Germany. This research site (*Merzalben*, 49° 16'N, 7° 48'E) consists of a number of permanent forest monitoring plots providing a large pool of in-situ reference measurements. The present study was carried out at the main monitoring plot, a mixed Beech Douglas fir stand, which is characterized by two zones of pure Beech and pure Douglas fir and one zone where these species evenly mix (see Table 1 for the basic stand characteristics). During a field campaign in August 2011 the pure zones were scanned with a total of 12 terrestrial laser scans (6 scans each).

Table 1: Stand characteristics of the research site *Merzalben*.

	DBH [cm]		Height [m]		Area [ha]	Stem density [n/ha]
Species	Mean	St. Dev.	Mean	St. Dev.	Total	Total
Beech	16.4	7.3	18.5	5.7	0.5	1032
Douglas fir	28.7	7.5	27.1	4	0.5	579

2.2 TLS scan collection and data processing

TLS scans were collected with a FARO Photon 120, a phase-shift laser scanner with 360° x 320° field of view and a minimum horizontal and vertical step size of 0.009°. The scanner has a maximum measurement speed of 976 kpt/sec and a beam divergence of 0.009°. The scanner settings used for this study were an angular resolution of 0.036° (point spacing of 6.28 mm/10 m) and a measurement speed of 122 kpt/sec. These settings resulted in a scan time of 6:49 min and file sizes of approximately 185 MB. In order to post-process the single scans into one merged scan, reference objects were placed in positions where they could be viewed from all single scan locations. Reference DBH (diameter at breast height) tape measurements were collected in April 2012. Tree locations were extracted from an existing database. In order to validate the stem volume estimates, stem diameter profiles (step size of 20 cm and 10 m height) were recorded in Mai 2012 for 6 Beech trees. The single scans were manually registered using FARO scene software to produce merged point clouds for each of the pure zones scanned. The single scan point clouds were also automatically converted to raster images containing the x, y,

z, range, intensity, azimuth and zenith information for each scan point.

2.3 Stem detection

Our approach to automated tree stem detection is based on range differences between neighboring pixels of a horizontal slice extracted from the range image of a single TLS scan. While single scan range images have been used for stem detection before (e.g. Forsman and Halme, 2005, van Leeuwen et al., 2011), the proposed approach is original and differs distinctly from these approaches. To minimize omission due to occlusion by branches and leaves our approach utilizes multiple slices. Scanner returns are assessed in 9 horizontal slices of 0.36° vertical thickness with the following zenithal offsets from the scanner height (= 0° zenith): -3.6°, -2.7°, -1.8°, -0.9°, 0°, 0.9°, 1.8°, 2.7°, 3.6°. Due to its lower stem density only 5 of these slices were used in the stem detection of the Douglas fir zone. Within each slice the range mean μ_r and variance σ^2_r are computed for each vertical scan line. If $\mu_r < \tau_{range}$ and $\sigma^2_r < \tau_{\sigma^2}$ then the scan line is considered to be a possible stem return. Consecutive stem returns with range differences below τ_{diff} and which cover a minimum azimuth span $\delta(\varphi)$ provide the primary evidence of a tree stem. Stem centre locations (i.e. their central azimuth) detected in each slice are then compared to determine if they correspond to a single tree stem or to different stems.

The parameter set used in the present study is: τ_{diff} = 10 cm, τ_{range} = 30 m (optional parameter), τ_{σ^2} = 0.0001, $\delta(\varphi)$ = 0.18°. Based on the scan resolution used in this study, an azimuth span of 0.18° corresponds to stem diameters of 3.14 cm, 6.28 cm, and 9.24 cm for ranges of 10 m, 20 m, and 30 m respectively. $\delta(\varphi)$ was determined based on the observation that less than 5% of the trees within τ_{range} have DBH below 10 cm.

2.5 Stem diameter and volume extraction

In order to assess the effects of the circle fitting on the extraction of stem diameters and volumes, three different algorithms were tested: A geometric algorithm implemented by J. R. Lemen (1991) and two algebraic algorithms (Pratt, 1987, and Taubin, 1991) as implemented by N. Chernov (2009). Geometric and algebraic fits are the two general circle fit approaches. While the geometric fit, which aims at minimizing the sum of the squares (SS) of the residuals, is considered to be accurate, a major concern is that the respective minimization algorithms have no closed solution and, thus, they usually require iterative and computationally intensive numeric schemes (Al-Sharadqah and Chernov, 2009). Algebraic fits on the other hand use an algebraic equation to represent a circle and corresponding fits are non-iterative and thus faster than Geometric fits (Pratt, 1987). One of the most popular algebraic fits is the Kåsa algorithm (Kåsa, 1976). However, it has been found that the accuracy of the Kåsa fit suffers in cases when the observed points do not represent complete circular arcs (Al-Sharadqah and Chernov, 2009). Thus, several modifications were developed to overcome this limitation, e.g. by Pratt (1987) and by Taubin (1991). To account for the incomplete circular representation of tree stems in single scan mode we tested the circle fit algorithms by Pratt and Taubin and in order to compare between the general circle fit approaches we included the geometric fit by Lemen (1991).

Our approach to an automatic stem diameter extraction can be applied to single and merged scan data. In a first step, scan points within a bounding box (60 cm x 60 cm) centred at the detected stem locations are extracted. Each tree defines its own set of stem points. The height minimum of each set representing the ground level is then determined. Since it has been found that the phase-shift FARO scans suffer from so-called ghost points below ground level (Bienert et al., 2006, Maas et al., 2008) which strongly affect the ground level and breast height determination, a histogram-based outlier detection and removal is implemented in our approach.

Scan returns classified as stem points are extracted for horizontal slices of 5 cm thickness and circle fitting is applied. Since the stem points usually include scan returns from branches and leaves which affect the circle fitting, we implemented a range-based test for these misclassified scan points. The importance of extracting meaningful information from TLS data is stressed by a number of authors (e.g. Aschoff et al., 2004, Maas et al., 2008, Litkey et al., 2008) and may be achieved by filtering and/or applying robust modeling techniques. To prevent gross circle fit errors we set an upper diameter threshold. The circle fitting is carried out for all slices from the local minimum to the maximum height. In a final step DBH is extracted as the stem diameter at the nominal height of 1.3 m as well as estimated based on a robust linear fit of the stem diameter profile. Stem volume is calculated based on the extracted and measured diameter profiles by summing the volumes calculated for each slice approximated as a conical frustum. Due to the increasing shadowing of stem parts within the canopy, stem volume calculation was restricted to stem heights of 10 m.

3. Results

3.1 Stem detection

The results reveal two important influences on the stem detection: The influence of the shadowing of stems by other vegetation components, and the influence of stem density. While the former has been well documented (e.g. Bienert et al., 2007, Litkey et al., 2008, Liang et al., 2009, Lovell et al., 2011), the effect of stem density has received relatively little attention in the TLS literature, despite the fact that stem density and the shadowing effect are directly related. The combined effects of shadowing and stem density can be particularly strong if stem detection is based on single TLS scans. This is reflected in the low detection rates based on using just one horizontal slice from the single scan range images (Tab. 2). Increasing the number of slices reduces the possibility of shadowing caused by branches and leaves and increases the detection rates to 94% and 96%. These values reduce to 85% and 84% if trees which are completely shadowed within the zenith range used for stem detection are accounted for. Comparing the results between the Beech and the Douglas fir zone demonstrates the influence of stem density on the detection rate: Using only one slice 12% more of the trees are detected in the lower density Douglas fir zone than in the higher density Beech zone. Using 5 slices 96% of the Douglas fir trees compared to 81% of the Beech trees are detected. This demonstrates the necessity to adjust the stem detection algorithm for stands of different stem density. However, since the algorithm is fast and efficient (~ 1–2 seconds per slice), the computational cost in increasing the number of slices is low, hence the initial number can be set relatively high. Concluding, overall detection rates and - equally important - false detection rates are excellent for both Beech and Douglas fir (Tab. 2).

Table 2: Stem detection statistics. Values are mean values of the 6 single scans. Occlusion relates to the degree of shadowing of stems by branches and leaves within the zenith range used for stem detection, i.e. occluded trees are still visible in the range image. By contrast all trees relates to visible and non-visible trees present within the range image.

N° of slices	Beech	Douglas fir
1	56%	68%
5	81%	96%
9	94%	
False detection rate	2%	1%
Detection rate unoccluded and partially occluded trees	94%	96%
Detection rate unoccluded and occluded trees	88%	86%
Detection rate all trees	85%	84%

3.2 DBH extraction

In order to test the influence of the scan mode and circle fitting on the stem diameter extraction, DBH was extracted from both single and merged scan data and based on three different circle fit algorithms and two different extraction methods (cf. section 2.5). DBH analysis is based on samples of 289 (Beech) and 219 trees (Douglas fir) for the single scans and 82 (Beech) and 67 trees (Douglas fir) for the merged scans. The results reveal three distinct aspects of the DBH extraction from phase-shift TLS data:

- The performance of the DBH extraction from the merged scans is superior to the extraction from the single scans.
- The performance of the DBH estimated based on a robust linear fit of the diameter profile is superior to the DBH extracted at the nominal breast height.
- The differences in the DBH estimates resulting from applying different circle fit algorithms are marginal.

The latter is reflected in the error histograms (Fig. 2). Single scan RMSE's are nevertheless higher for the algebraic-based algorithms Pratt and Taubin than for the parametric-based Lemen algorithm – this is true for both Beech and Douglas fir and to a lesser degree for both DBH extraction methods (Tab. 3), which consolidates the slightly better performance of the Lemen algorithm which also proved to be more robust in case of noisy data. These differences, however, diminish for the merged scan RMSE's (Tab. 3), which demonstrates that the influence of the circle fitting can be neglected for merged scans. Results also show that the single scan DBH extracted at the nominal breast height differ significantly from those estimated based on a linear fit of the diameter profile (Fig. 2), which can be attributed to the fact that the latter is less prone to errors in the determination of the ground level and thus the exact breast height as well as to errors in the circle fitting due to outliers. The larger RMSE differences between the two extraction methods compared to those between the three circle fit algorithms indicate that a robust modeling and/or removal of outliers is more important for DBH extraction from single TLS scans than the choice of the circle fit algorithm.

Table 3: Stem diameter statistics. DBH (BH) = DBH extracted at the nominal breast height, DBH (Lfit) = DBH estimated based on fitting a line to the diameter profile.

		Single scans RMSE [cm]		Merged scans RMSE [cm]	
Plot	Algorithm	DBH (BH)	DBH (Lfit)	DBH (BH)	DBH (Lfit)
Beech	Lemen	1.58	1.39	0. 66	0. 67
	Pratt	2.20	1.49	0. 90	0. 64
	Taubin	2.43	1.57	0. 79	0. 66
Douglas fir	Lemen	1.47	1.65	0. 97	1.06
	Pratt	2.04	1.68	1.21	1.15
	Taubin	2.02	1.74	1.20	1.14

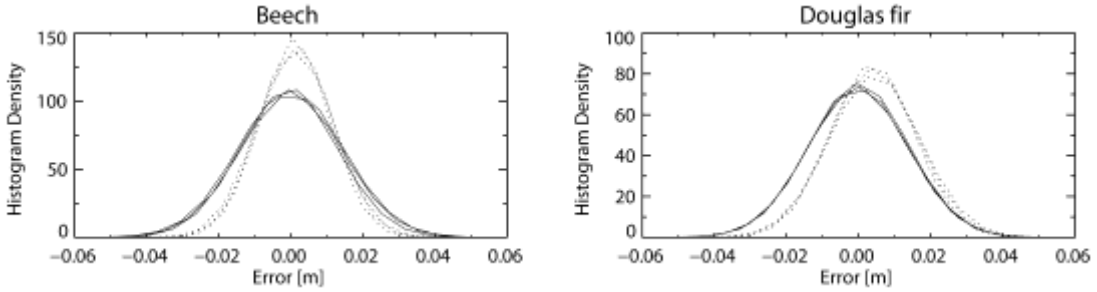


Figure 1: Distribution of the single scan DBH estimation errors for the different DBH extraction methods and circle fit algorithms. Dotted lines = Circle fit errors based on the linear fit DBH extraction. Solid lines = Circle fit errors based on the direct DBH extraction.

Contrary to the single scans, differences in RMSE between the two extraction methods are smaller for the merged scans (Tab. 3), which suggests a reduced influence of outliers. The biggest difference compared to the single scan DBH extraction, however, are the considerable lower RMSE's irrespective of the extraction method, tree species and circle fit algorithms (Tab. 3). The improved accuracy can be attributed to a more stable and accurate circle fit which results from the full coverage of stem cross-sections and the reduced influence of outliers due to this improved coverage. As already mentioned, this is demonstrated by the reduced influence of the DBH extraction methods and circle fit algorithms for the merged compared to the single scans (Tab. 3).

3.3 Stem volume extraction

The estimation of stem volume is a main objective in forest inventory, in particular from a commercial point of view (Bienert et al., 2007). By contrast to traditional methods, TLS potentially allows for a direct measurement of stem volume and therefore could be a valuable tool for forest inventory. One of the aims of this study was to assess the differences in the stem volume estimates based on single and merged scan data and based on different circle fit algorithms. While stem volumes extracted from the merged scans show an excellent agreement, with mean deviations ranging from -1% to 6%, stem volumes extracted from the single scans vary from -18% to 25% (Tab. 4). The higher variability mainly results from the inferior stem diameter extraction based on the single TLS scans (cf. section 3.2). However, the results also demonstrate that it is possible to achieve good stem volume estimates from single scans with errors as low as 5% (Tab. 4), however this requires a careful and potentially time-consuming scanner positioning. As with the stem detection the results reveal the main limitation of the single scan mode which is the shadowing by stems, branches and leaves. Due to this effect only 1/3 to 2/3 of the single scans could effectively be used to extract stem profiles and volumes in this study (Tab. 4).

Table 4: TLS stem volume deviations from the reference volume [%]. Deviations are mean values of the three different circle fit algorithms. The dash indicates either an incomplete stem profile due to partial occlusion or non-visibility to the scanner due to complete shadowing.

Tree	Scan 1	Scan 2	Scan 3	Scan 4	Scan 5	Scan 6	Merged Scans
542	-	16	-	12	-	-14	1
543	-6	5	17	10	-	-	-1
544	-	-	25	20	-	-	3
546	-	-	-9	5	15	-	1
810	-	-	17	-13	18	-18	2
823	23	-	-	-	-	15	6

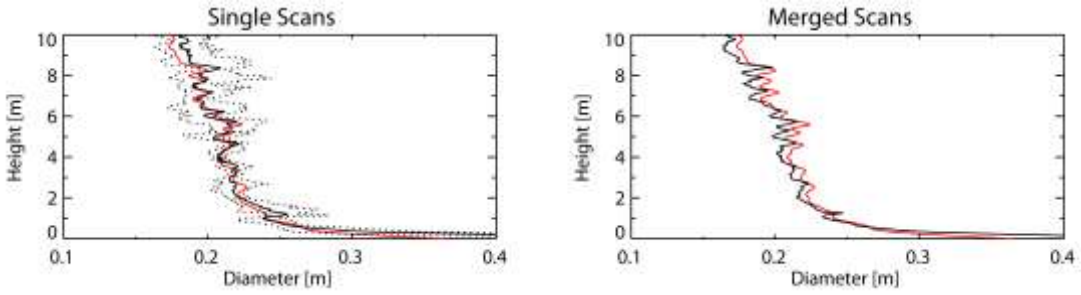


Figure 2: Comparison of reference and stem profiles derived from the single and merged TLS scans. Red line = reference profile, dotted lines = min. and max. of single scan profiles, solid line = stem diameter profiles derived from the single (mean) and merged scans.

4. Discussion

Research in recent years has proven that terrestrial laser scanning (TLS) is a valuable tool to assess vegetation structure, in particular with regard to traditional forest inventory parameters (e.g. Maas et al., 2008). However, there are number of factors which influence the information extraction from TLS data and which necessitate further research, in particular if TLS is to be routinely used in forest inventory. The present study therefore focused on assessing the influence of circle fitting and scan mode on the automatic extraction of stem diameter and volume. To assess the effect of stem density, scans were collected for two different forest plots (579 compared to 1032 stems/ha). We propose a fully automated and efficient approach to stem detection, stem diameter and volume extraction. Stem detection is based on range differences between neighbouring pixels of horizontal slices extracted from the range images of single TLS scans. Mean stem detection rates range from 94% to 96% for trees with an unobstructed field of view to the scanner and from 84% to 85% for all trees within a distance of 20 m from the scanner. Equally important, false detection rates are low (1% to 2 %). Analysis revealed that both stem density and number of slices used in the stem detection algorithm have a significant effect on detection rates. This can be attributed to the main limitation of the single scan approach, the shadowing of stems by other vegetation components. Shadowing depends on the stand structure which itself is influenced by a number of factors (e.g. stem density, tree species composition and age, canopy layering, degree of undergrowth, and forest management practices such as thinning). The shadowing effects could be reduced by using merged scans which have the advantage of an improved 3D coverage (e.g. Bienert et al., 2006), but may not lend themselves easily to stem detection approaches as simple and efficient as those based on single TLS scans.

The proposed extraction of stem diameter and volume is based on circle fitting to stem point cloud slices of single and merged scans. Additionally, three different circle fit algorithms were tested. DBH is determined based on extracting the diameter at the nominal breast height as well as based on a robust linear fit of the diameter profile. Both methods yield DBH which agree closely with reference DBH. The linear fit approach, however, proved to be more robust for the single scans (RMSE's range from 1.39–1.74 cm compared to 1.47–2.43 cm for the direct extraction method). With regard to the different circle fit algorithms, the Lemen (1991) method showed the best overall performance and was also found to be more robust in case of noisy data. This is reflected in lower RMSE's for the direct DBH extraction which is prone to errors in the ground level determination. Compared to the single scans, the DBH extraction from the merged scan data proved to be superior with significant lower RMSE's (0.66–1.21 cm). The improved accuracy is due to a more stable circle fitting as a result of the full coverage of stem

cross-sections, which also has the advantage of reducing the influence of outliers as reflected in the smaller differences between the DBH estimates from the different circle fit algorithms and DBH extraction methods (Tab. 3). The superior performance of the diameter extraction from the merged scans is also reflected in the corresponding stem volume estimates which exhibit only small deviations from the reference volume (-1% to 6%) compared to the strongly varying single scan estimates (-18% to 25%).

This study demonstrates that the scan mode has a significant effect on the stem diameter and volume extraction from TLS data, i.e. estimation errors based on merged scans are lower compared to those based on single scans. By contrast the effect of using different circle fit algorithms is marginal and can be neglected. However, since the factors influencing the parameter retrieval from TLS data are interconnected, further research into their combined effects is required. This will be addressed in a follow-up study. While merged scan data performs superior in terms of the parameter estimation, it suffers from a tedious and time-consuming measurement set-up. This indicates an important consideration for the integration of TLS into forest inventory: the optimization of the sampling design (number and positioning of scans, scan resolution, etc.) to ensure sampling efficiency. This may also include an optimization with respect to the basic inventory parameters and to the variability in stand structure. Despite these issues TLS has proven to be a valuable tool for the retrieval of structural forest parameters and it is to be expected that TLS will become a standard in forest inventory.

Acknowledgements

We are very grateful to the many colleagues and students who participated in the field campaign and supported this study. The authors would in particular like to thank R. Bögelein for help with the acquisition of the stem diameter profiles. This research was supported within the framework of the EnMAP project (contract No. 50EE0946-50) by the German Aerospace Center (DLR) and the Federal Ministry of Economics and Technology.

References

Al-Sharadqah, A. and Chernov, N., 2009. Error analysis for circle fitting algorithms. Electronic Journal of Statistics, 3: 886-911.

Aschoff, T. and Spiecker, H., 2004. Algorithms for the automatic detection of trees in laser scanner data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVI-8/W2: 66-70.

Bienert, A., Scheller, S., Keane, E., Mohan, F. and Nugent, C., 2007. Tree detection and diameter estimations by analysis of forest terrestrial laserscanner point clouds. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVI-3/W52

Bienert, A., Scheller, S., Keane, E., Mullooly, G. and Mohan, F., 2006. Application of terrestrial laser scanners for the determination of forest inventory parameters. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVI part 5.

Chernov, N., 2009. Circle fit (Pratt method), MATLAB CENTRAL. http://www.mathworks.com/matlabcentral/fileexchange/22643-circle-fit-pratt-method (accessed 26.07.2012)

Chernov, N., 2009. Circle Fit (Taubin method), MATLAB CENTRAL. http://www.mathworks.com/matlabcentral/fileexchange/22678 (accessed 26.07.2012)

Forsman, P. and Halme, A., 2005. 3-D mapping of natural environments with trees by means of mobile perception. Robotics, IEEE Transactions on, 21(3): 482-490.

Holopainen, M. et al., 2011. Biomass estimation of individual trees using stem and crown

diameter TLS measurements, ISPRS Workshop Laser Scanning Calgary, Canada.

Jupp, D.L.B. et al., 2009. Estimating forest LAI profiles and structural parameters using a ground-based laser called Echidna®. Tree Physiology, 29(2): 171-181.

Kåsa, I., 1976. A curve fitting procedure and its error analyis. IEEE Trans. Inst. Meas., 25: 8-14.

Lemen, J.R., 1991. Fit circle. Astronomy Department, University of Washington. http://www.astro.washington.edu/docs/idl/cgi-bin/getpro/library32.html?FIT_CIRCLE (accessed 26.07.2012).

Liang, X. et al., 2009. Automatic stem location mapping using TLS for plot-wise forest inventory, SilviLaser, College Station, Texas, USA.

Litkey, P. et al., 2008. Single-scan TLS methods for forest parameter retrieval, SilviLaser, Edinburgh, UK.

Lovell, J.L., Jupp, D.L.B., Newnham, G.J. and Culvenor, D.S., 2011. Measuring tree stem diameters using intensity profiles from ground-based scanning lidar from a fixed viewpoint. ISPRS Journal of Photogrammetry and Remote Sensing, 66(1): 46-55.

Maas, H. G., Bienert, A., Scheller, S. and Keane, E., 2008. Automatic forest inventory parameter determination from terrestrial laser scanner data. International Journal of Remote Sensing, 29(5): 1579-1593.

Pratt, V., 1987. Direct Least-Squares Fitting of Algebraic Surfaces. Computer Graphics, 21(4): 145-152.

Simonse, M., Aschoff, T., Spiecker, H. and Thies, M., 2003. Automatic determination of forest inventory parameters using terrestrial laserscanning, ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests, Umea, Sweden.

Tansey, K., Selmes, N., Anstee, A., Tate, N.J. and Denniss, A., 2009. Estimating tree and stand variables in a Corsican Pine woodland from terrestrial laser scanner data. International Journal of Remote Sensing, 30(19): 5195-5209.

Thies, M. and Spiecker, H., 2004. Evaluation and future prospects of terrestrial laserscanning for standardized forest inventories. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36: 192-197.

Pratt, V., 1987. Direct least-squares fitting of algebraic surfaces. Computer Graphics, 21(4): 145-152.

Van Leeuwen, M., Coops, N.C., Newnham, G.J., Hilker, T., Culvenor, D.S., Wulder, M.A., 2011. Stem detection and measuring DBH using terrestrial laser scanning. SilviLaser 2011, Oct. 16-20, 2011 – Hobart, Australia.

Yao, T., Yang, X.Y., Zhao, F., Wang, Z.S., Zhang, Q.L., Jupp, D., Lovell, J., Culvenor, D., Newnham, G., Ni-Meister, W., Schaaf, C., Woodcock, C., Wang, J.D., Li, X.W., & Strahler, A. (2011). Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. Remote Sensing of Environment, 115, 2965-2974

Modelling variation in Pinus radiata stem velocity from area- and crown- based LiDAR metrics

David Pont[1], Michael S. Watt[2], Thomas Adams[3], Hamish Marshall[4], John Lee[5], David Crawley[6], Pete Watt[7]

[1]Scion, Rotorua, New Zealand, david.pont@scionresearch.com
[2]Scion, Christchurch, New Zealand, michael.watt@scionresearch.com
[3]Metservice, Wellington, New Zealand, thomas.adams@metservice.com
[4]Interpine Forestry Ltd, Rotorua, New Zealand, Hamish.marshall@interpine.co.nz
[5]Scion, Rotorua, New Zealand, john.lee@scionresearch.com
[6]PF Olsen, Rotorua, New Zealand, david.crawley@pfolsen.com
[7]Indufor Asia-Pacific Ltd, Auckland, New Zealand, pete.watt@indufor-ap.com

Paper Number: ### SL2012-081

2. Introduction

Assessment of standing forest resource is essential to effective forest management. Currently assessment in New Zealand is largely based on ground measurements. LiDAR is a remote sensing technology rapidly being adopted for applications in forest management. Aerial LiDAR scanning (ALS) is already used in the New Zealand national Carbon assessment system (LUCAS, Stephens *et al*., 2012) and is used operationally for forest management in Australia by Forestry Tasmania (Turner *et al*., 2011). Operational use of ALS in forestry is largely based on area-based metrics (often referred to as grid metrics), where statistical properties of the spatial distribution of LiDAR points are generated over a rectilinear grid. The statistical grid metrics are then used in models to estimate variables of interest, such as: mean top height, basal area, stem volume and carbon. Because the metric data used in modelling is on a grid (say 20 by 20m) the variables of interest are thus able to be mapped as raster layers at this resolution. The variables can also be averaged to obtain stand or stratum level estimates as conventional ground inventory would provide.

A key premise of our research project has been that measurement of crown characteristics such as crown size and asymmetry might deliver more detailed information about a resource, perhaps in terms of piece size, log grade or wood quality. Methodology to segment individual trees has been applied to LiDAR data. Further processing of the segmented LiDAR data allows the calculation of a number of crown-based metrics for each tree. Chen *et al*. (2007) and Jung *et al*. (2011) used individual tree crown metrics derived from LiDAR to estimate tree basal area, volume and biomass, but did not investigate wood properties. Maltamo et al. (2009) successfully used a mixture of area- and crown- based LiDAR metrics to estimate: DBH, height, crown height, height of lowest branch, volume, and proportion of sawlogs for Scots pine in Finland. In a recent review paper the benefits and potential of assessing wood quality from LiDAR were clearly described (van Leeuwen et al., 2011). This paper covers the use of area- and crown- based LiDAR metrics to estimate standing tree outerwood acoustic velocity (V), which is known to be correlated with stiffness – an important wood quality measure.

2. Method

2.1 Data collection

2.1.1 Study area

The data for this study was obtained from a forest located in Eastern Bay of Plenty on steep country with an elevation range of approximately 150 to 300 metres above sea level. The forest is a first rotation planted onto cleared native forest.

2.1.2 Plot establishment

A standard pre-harvest inventory was installed using the stand/harvesting boundaries as inventory population. In total 172 plots were installed in the stands where the trees were still standing at the time of inventory. These plots represented a sample area of 217.8 hectares. The circular plots were laid out within each population on a systematic grid as is normal practice in New Zealand forest inventory. The plot size varied from population to population so that a target of approximately of 20 trees per plot could be obtained. The exact location of the plots were measured using a high grade GPS and post processing differential correction carried out. Within each plot all trees were measured for diameter with a subset of heights being measured. The field data was collected in the early part of 2011.

2.1.3 Outerwood velocity

Outerwood acoustic velocity (V) was measured in 35 of the inventory plots. Plots were selected to cover the range in stem slenderness and stocking present throughout the forest. Where possible, measurements were taken from at least 20 trees within the plot.

Velocity measurements were taken using the ST300 tool. Due the steepness of the site, these were taken from the uphill side of the tree and centred where possible on breast height (1.4m). Two measurements were taken on each stem, with probes placed approximately 1m apart and avoiding paths that would include any large branch stubs or obvious malformations. Measurement paths were located as close as possible to 180 degrees apart.

2.2 Predictive variables included in the modelling

2.2.1 LiDAR dataset

The study site was flown for LiDAR data and aerial imagery by New Zealand Aerial Mapping between 24 May and 1 June 2011, using NZ Aerial Mapping's Optech ALTM 3100EA LiDAR system (05SEN178) and Trimble AIC medium format digital camera. The LiDAR data was collected at a minimum of 2pts per m^2 on open ground. The raw LiDAR data was processed by the supplier into LAS format, georeferenced into the New Zealand Transverse Mercator (NZTM) coordinate system.

2.2.2 Area-based metrics

The software package FUSION (McGaughey *et al*. 2004) was used for many of the LiDAR processing tasks. Plot locations, measured using high grade differentially corrected GPS, were used to extract sections of LiDAR data relating to each plot, assuming a circular plot centred on the GPS location. Fusion (ClipData) was used for this, as well as generating the area-based LiDAR height metrics used in analysis (CloudMetrics). The height metrics were generated using a manually corrected Digital Terrain Model (DTM) supplied by the data provider. All returns within 0.5 m of the ground were eliminated to remove the effects of understory.

2.2.3 Crown-based metrics

Fusion was used to generate a canopy height model (CHM) from the LiDAR (CanopyModel) and converted to a TIF image (DTM2TIF). The CHM TIF image was processed with a watershed segmentation process calibrated by comparing numbers of trees in ground plots (the pre-harvest inventory plots in this case) with numbers of trees identified in the same location on the CHM, minimising over- or under segmentation. The output of the segmentation process is a set of boundaries around individual trees. The boundaries are adjusted to exclude areas of ground or low understory adjacent to some trees, giving more accurate crown boundaries. The boundary and crown three-dimensional shape, as represented in the CHM image, are analysed to generate a number of crown metrics for each tree. Plot locations are used to average crown metrics for all trees falling within each plot area and a file of plot-level crown metrics created. The plot-level crown-based metrics were then combined with the ground plot data and area-based LiDAR plot metrics for analysis.

2.2.4 Grid metrics for model estimation

In order to extrapolate the relationship established at the sampled plot locations to the whole forest, FUSION was used to generate LiDAR grid metrics for the forest on a 20m x 20m (0.04ha) grid (GridMetrics). This grid was converted to ASCII text (CSV2Grid) and combined with individual crown metrics averaged onto the same rectilinear grid. The combined area- and crown- based metric data was then clipped with stand area shapefiles, and the predictive equation for V applied in the Python numerical programming environment. This produced a spatial map of V within the stand areas.

2.3 Analysis

Models used to predict outerwood velocity were generated using SAS (SAS-Institute-Inc 2000), using a general linear model (PROC GLM). Variables were introduced sequentially into each model starting with the variable that exhibited the strongest correlation, until further additions were either: not significant; not physiologically reasonable; or did not markedly improve model precision.

Variable selection was undertaken manually, one variable at a time, and plots of residuals were examined prior to variable addition to ensure that the variable was included in the model using the least biased functional form.

3. Results

3.1 Crown-based metrics

Figure 1 shows the CHM image input to the segmentation process and Figure 2 shows the crown boundaries produced for stand area 1.

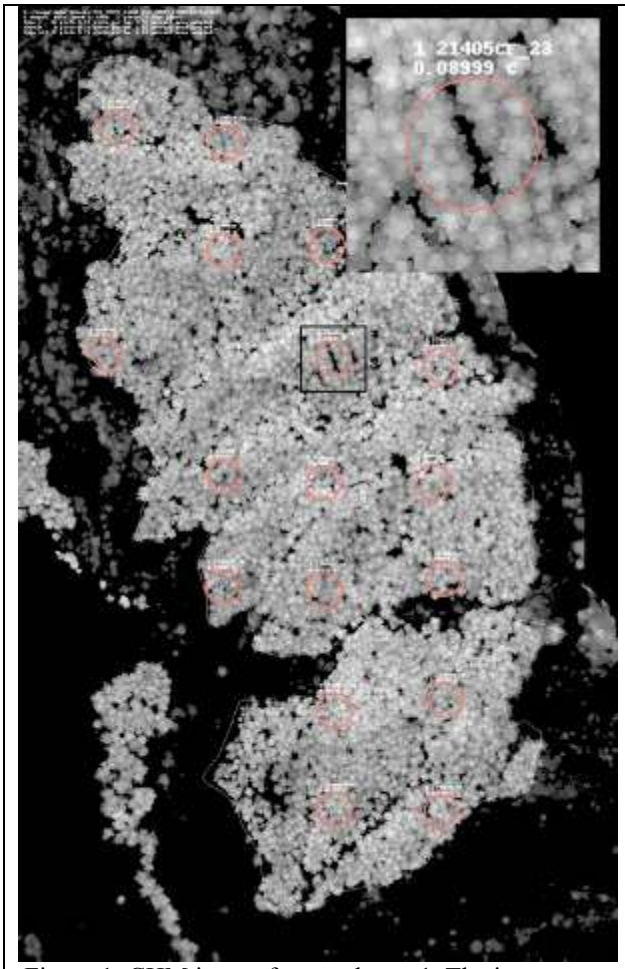


Figure 1: CHM image for stand area 1. The inset at top right shows the area in the black box magnified. Plot boundaries are outlined in light red.

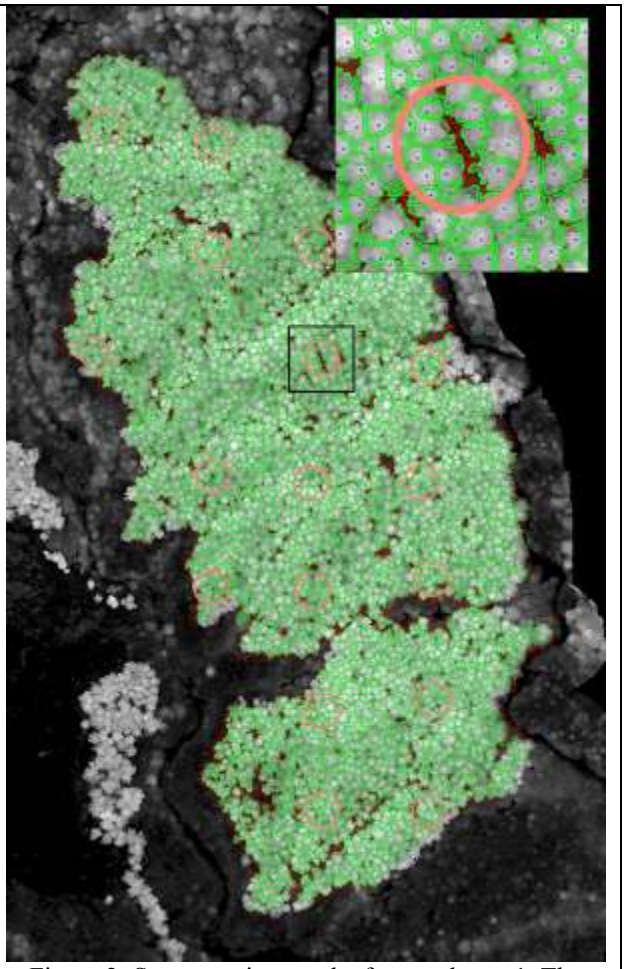


Figure 2: Segmentation results for stand area 1. The inset at top right shows the area in the black box magnified. Plot boundaries are outlined in light red. Crown boundaries are drawn in green. Tree tops are marked with a blue dot.

A total of 4671 trees were identified in stand area 1 and a number of metrics calculated for each crown. Crown metrics included: projected area, surface area, volume, perimeter, radius, compactness and aspect ratio.

The best statistical model that included only area-based LiDAR metrics explained 27% of the variation in V. In this model the standard deviation in LiDAR heights was included as a positive linear term. The model had an RMSE of 0.13 km s-1 and was described as:

$$V = 3.63 + 0.00671 H_{SD} \quad (1)$$

The overall model was significant (P=0.0018) as was H_{SD}, standard deviation of height, (P<0.0018). A plot of model predictions against actual velocity showed little apparent bias (Fig. 3) although there was one outlier with low velocity.

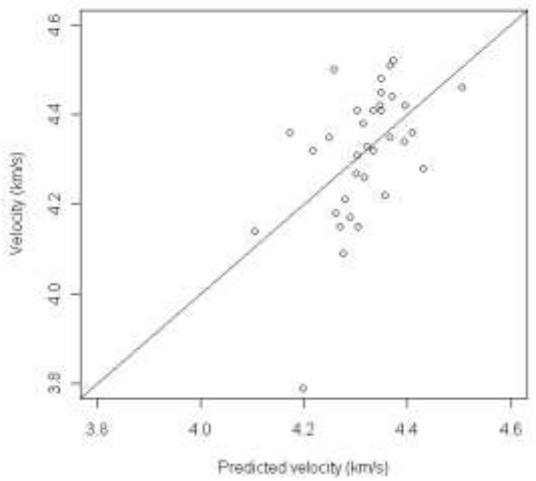


Figure 3. Relationship between predicted and actual velocity developed from model with LiDAR only metrics. For reference the 1:1 line is shown as a dashed line.

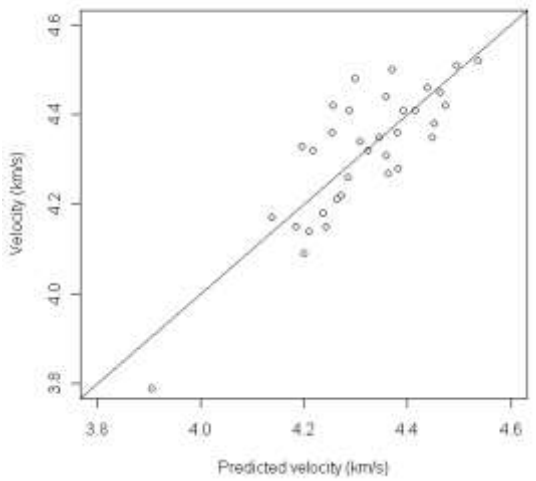


Figure 4. Relationship between predicted and actual velocity developed using area- and crown-based LiDAR metrics. For reference the 1:1 line is shown as a dashed line.

A model using a mixture of area-based and crown metrics was more precise explaining 69% of the variation in V. In this model the coefficient of variation for LiDAR heights and crown aspect ratio were included as positive linear terms while the crown polygon area was included as a negative linear term.

The model had an RMSE of 0.09 km s-1 and was described as:

$$V = 3.90 - 0.01004C_{PA} + 0.902H_{CV} + 0.291C_{AR} \quad (2)$$

The overall model was significant as were all variables ($P<0.05$). A plot of predictions against actual velocity was relatively unbiased (Fig. 4).

3.2 Spatial description of velocity

Spatial variation in V using the model described by Equation 2, is shown in Fig. 5. Velocity was variable within stands and there were slight but observable differences between some stands.

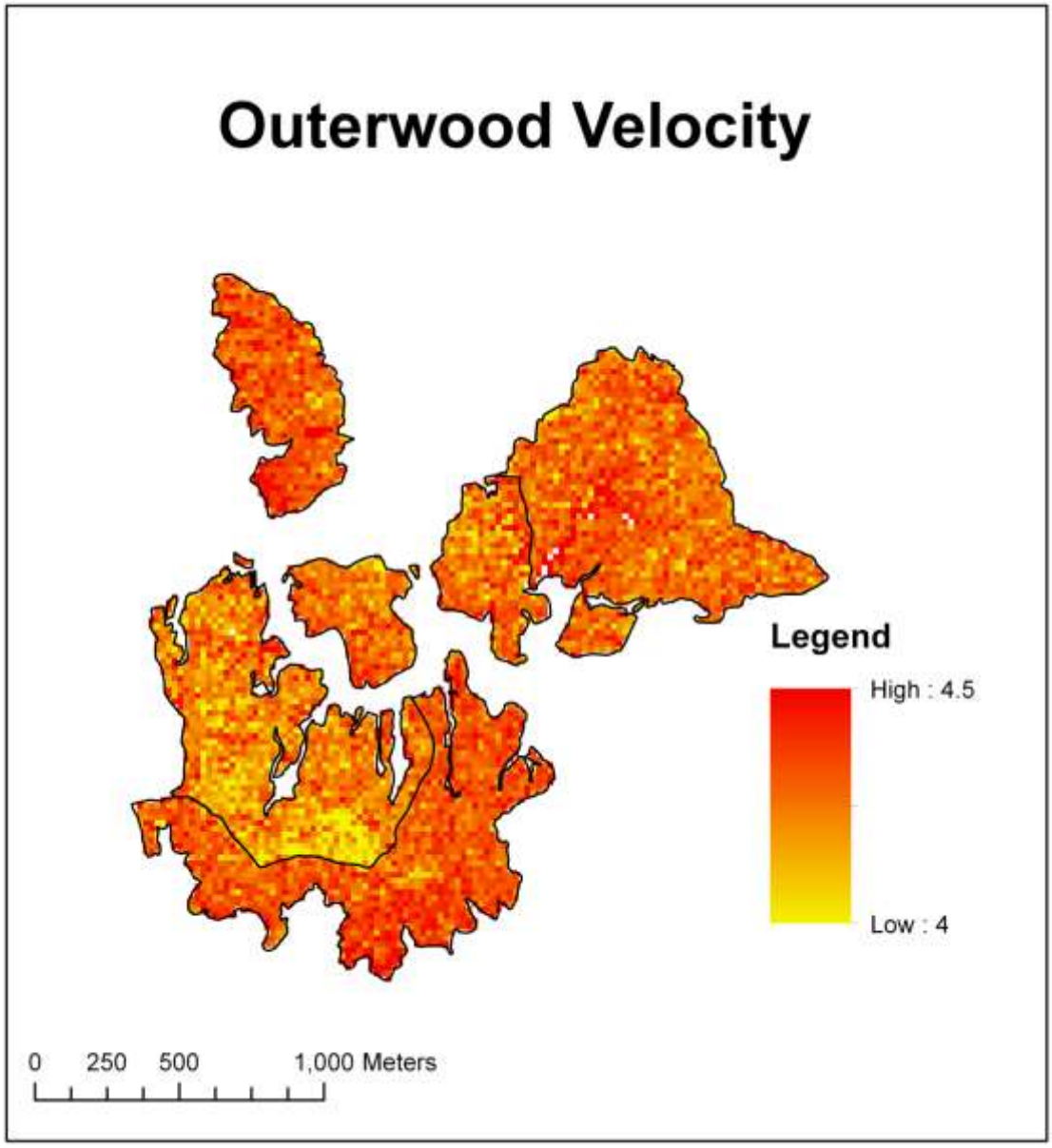


Figure 5. Spatial distribution of velocity predicted with the model using crown- and area- based LiDAR metrics (see Eqn. 2).

4. Discussion

Close examination of velocity estimates and crown segmentation showed low velocity often occurred in areas of trees with large crowns. This observation is supported by the main variable in the model (Eqn. 2) being crown polygon area, the projected area of the crown, as a negative linear term. Trees with large crowns often occurred at stand edges or adjacent to internal stand gaps. A clear change from high to low velocity is predicted at the boundary between two adjacent stands which lies on a ridge oriented east-west in the south west corner of the mapped area (Fig. 5). Examination of the canopy images at this boundary showed larger crown size north of the boundary, associated with lower stocking (see Fig. 6).

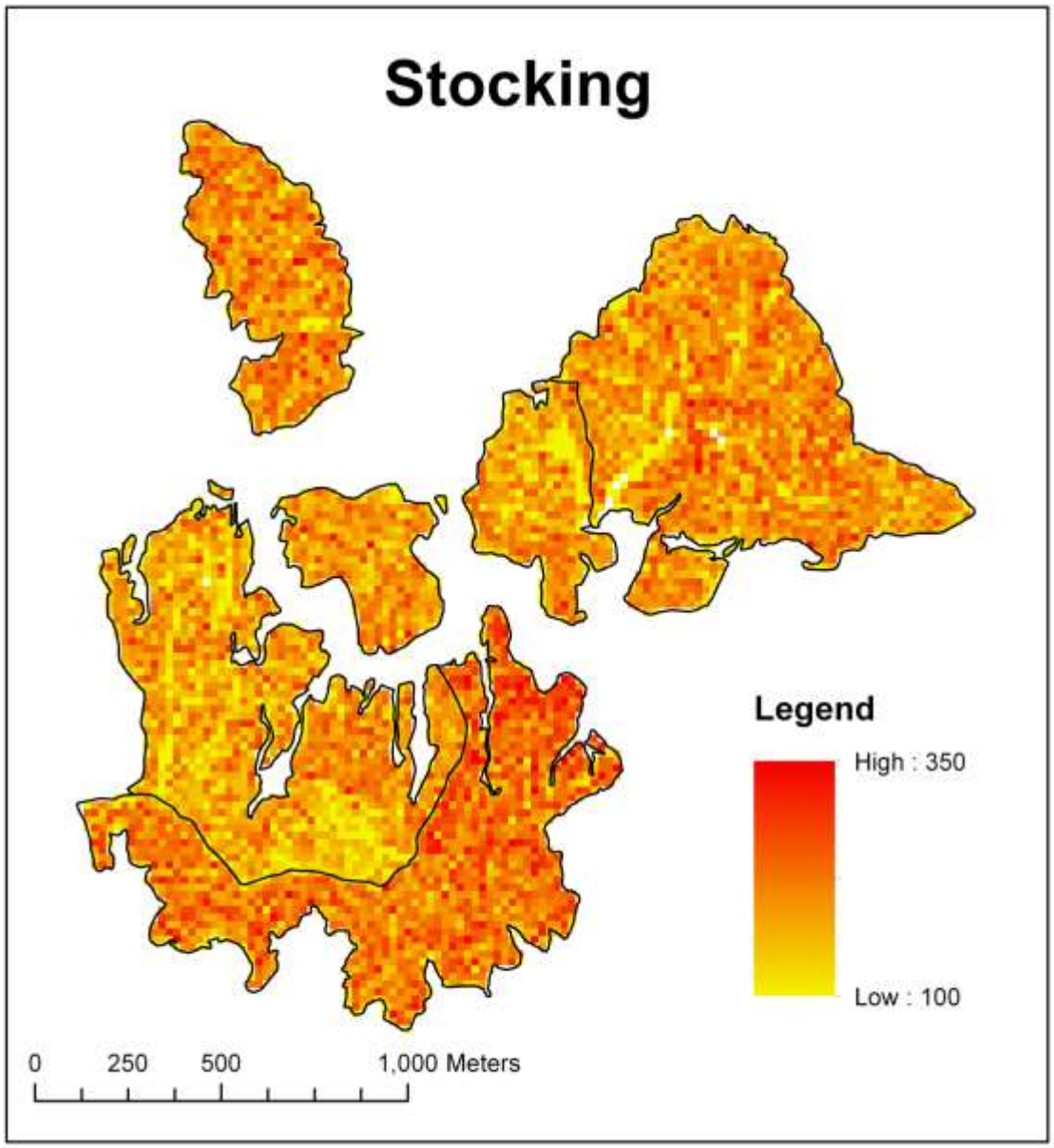


Figure 6. Spatial distribution of stocking.

The contribution of the variable H_{SD} to the first model (Eqn. 1), and the variables H_{CV} (coefficient of variation for LiDAR heights) and C_{AR} (crown aspect ratio) to the second model (Eqn. 2) are not so clear. Stephens *et al*. (2012) note that H_{CV} is often found useful for estimating basal area, volume and biomass, high H_{CV} associated with sparse open canopies, low H_{CV} with dense closed canopies. H_{CV} is H_{SD} (standard deviation of heights) divided by mean height, effectively a standardised form of H_{SD}. The two variables are similar; both being measures of the variability of height and appearing as positive linear terms in the two models. The implication is areas of sparse open canopies (high H_{CV} and H_{SD}) are associated with high V. It is not obvious how to reconcile this with the earlier observation that areas with small crowns (low C_{PA}) and low stocking are associated with high V.

The data analysed covered a relatively small range of V. The relationship between crown area (C_{PA}) and V is plausible. Interestingly on this dataset crown area was a better predictor than slenderness, a variable known to correlate well with V. Further analyses with data from sites with a broader range of V will be necessary to determine the ability to use area- and crown- based metrics in robust models of V.

Stiffness is widely reported to correlate well with V. Stiffness is a key wood quality variable, critical to determining yields of structural log and timber grades. The current methodology to generate crown-based LiDAR metrics relies on ground plots, used to calibrate a tree segmentation process. Area-based LiDAR metrics can be generated directly from the LiDAR point cloud, with no reliance on ground data. It remains to be established if the substantial improvement in prediction of V from using crown-based LiDAR metrics warrants the extra cost of establishing ground plots. If ground plots are already required as part of a complete inventory plan, then crown-based metrics can be made available with a minimal extra cost for the post-processing of the LiDAR.

Knowledge of the spatial variation in V, represented as raster maps, could be used in harvest planning. Better understanding of the factors determining V could be used to inform silviculture. If crown size, influenced strongly by stocking, is a key driver as indicated in this preliminary study, stocking could be managed to increase V.

Acknowledgements

This project was funded within the Radiata theme of Future Forests Research Ltd. We thank PF Olsen for provision of the LiDAR and PHI data; Interpine for carrying out the PHI inventory; John Lee and Tom Adams for collection of the wood quality data.

References

Chen, Q., Gong, P., Baldocchi, D., Tian, Y.Q., 2007. Estimating Basal Area and Stem Volume for Individual Trees from Lidar Data. *Photogrammetric Engineering & Remote Sensing*, Vol. 73, No. 12, 1355–1365.

Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P., Tokola, T. 2009. Predicting Tree Attributes and Quality Characteristics of Scots Pine Using Airborne Laser Scanning Data. *Silva Fennica* 43(3): 507–521.

McGaughey, R.J., Carson, W.W., Reutebuch,S.E., Andersen, HE., 2004. Direct measurement of individual tree characteristics from LIDAR data. *Proceedings of the Annual ASPRS Conference*, Denver, May 23-28, 2004. American Society of Photogrammetry and Remote Sensing, Bethesda, MD.

Stephens, P.R., Watt, P.J., Loubster, D., Haywood, A., Kimberley, M.O. 2012. Estimation of carbon stocks in New Zealand planted forests using airborne LiDAR. *Analysis*, XXXVI, 389-394.

Turner, R., Goodwin, N., Friend, J., Mannes, D., Rombouts, J., Haywood, A. 2011. A national review of airborne LiDAR application in Australian forest agencies. *Proceedings: Silvilaser 2011, 11th International Conference on LiDAR "Applications for Assessing Forest Ecosystems"*, 16-20 October, University of Tasmania.

van Leeuwen, M., Hilker, T., Coops, N.C., Frazer, G.W., Wulder, M.A., Newnham, G.J., Culvenor, D.S. 2011. Assessment of standing wood and fiber quality using ground and airborne laser scanning: A review. *Forest Ecology and Management*, 261(9): 1467-1478.

INDIVIDUAL TREES SPECIES CLASSIFICATION USING RELATIVE CALIBRATED FULL-WAVEFORM LIDAR DATA

Xu GuangCai[1], Pang Yong*[1], Li Zengyuana[1], Zhao Dan[1], Liu Luxia[1]

[1] The research institute of forest resources information technique, Chinese Academy of Forestry. Email: xuguangcaiyuan@126.com
caf.pang@gmail.com
zy@caf.ac.cn
zhaodan@irsa.ac.cn
liuluxiaok@126.com

Paper Number: SL2012-084

1. Introduction

The emergence of small-footprint full-waveform LIDAR provides a new train of thought for land-cover classification. The width, amplitude and intensity are the most important physical properties exploring from full-waveform data (Wagner et al., 2008). Persson et al. (2005) and Wagner et al. (2008) showed that the ground points tend to have lower echo widths and higher intensity compared to vegetation points. Therefore, low vegetation points can be identified and excluded before filtering terrain points solely based on the echo width and the backscatter cross section (Briese et al., 2008). And combining echo width and geometry information lead to a higher quality of the derived terrain model (Doneus et al., 2006; Ullrich et al., 2007). Forest areas are the main application of full-waveform LIDAR data since forest canopy tends to produce numerous multiple returns between the first and last pulse (Mallet et al., 2009). Hofle et al. (2008) indicate that mean echo width is applicable to separate larch from deciduous trees. Hollaus et al. (2009) use the standard deviation of the echo width per crown to further separate spruce from larch. Reitberger et al. (2007) demonstrated that the mean echo amplitude of stem points can assist the discrimination between coniferous and deciduous trees. Pulse width and tree geometry information were applied to single trees classification of coniferous and deciduous and lead in the best case to an overall accuracy of 85% in a leaf-on situation and 96% in a leaf-off situation (Reitberger et al., 2008). Neuenschwander et al. (2009) extracted 10 metrics from full-waveform data and selected the most important parameters for discriminating different tree types and densities. Johannes and Barbara (2011) utilized parameters extracted from full-waveform Lidar to classify six tree species with an overall accuracy of 57%, 78% to four main species and 91% to conifers and broadleaved trees.

Calibration is the necessary step in quantitative analysis of remote sensing and it is the foundation of a classification method for large area application. Small-footprint airborne LIDAR data is usually captured by multiple overlapped flights to cover a large area and many factors, such as emission pulse frequency, divergence angle, pulse width, scan angle, atmosphere, flight speed, distance, terrain, etc. may be having some changes in the different flights. So there may be some problems of classify different objects using the information extracted from full-waveform data directly without any calibration. Although some scholars have been carried out studies of LIDAR data calibration, only few initial studies combine the LIDAR data calibration and object classification especially in forestry application. Therefore, the objective of this paper is to use the calibrated Gaussian decomposition

results of full-waveform data to classify single trees species and evaluate the impact of the relative radiometric calibration of full-waveform LIDAR data on tree species identification.

2. Material

2.1 study area

The study area is located in the LingShui National Nature Reserve, Yichun City, Heilongjiang Province. The center geographical coordinates are about 128°53′20″E and 47°10′50″N, with an elevation ranging from 280 m to 707 m above the mean seal level and a slope of less than 40°. More than 98% percent of this area is covered by forest. Red Pine (Pinus koraiensis), Koyama Spruce (*Picea koraiensis*), Dahurian Larch (*Larix gmelinii*) and Fir (*Abies nephrolepis*) are the main coniferous species. White birch (*Betula platyphylla Suk*), Ribbed birch (*Betulla Costata*), Manchurian linden (*Tilia Mandschurica*), Elm (*Ulmus Laciniata*) and Mongolian oak (*Quercus Mongolica*) are the dominant deciduous species.

A 300m*300m square area was chosen from the LiangShui natural reserve, which includes almost all the main tree species of the flight LIDAR data (see Figure 1). For each tree in the plot, the species was recorded and the diameter at breast-height greater than 2 cm was measured including height, two direction crown diameters (east-west, south-north) and tree graduation. The precise position of individual trees was acquired by Differentiate Global Positioning System (DGPS) and Electronic Total Station.

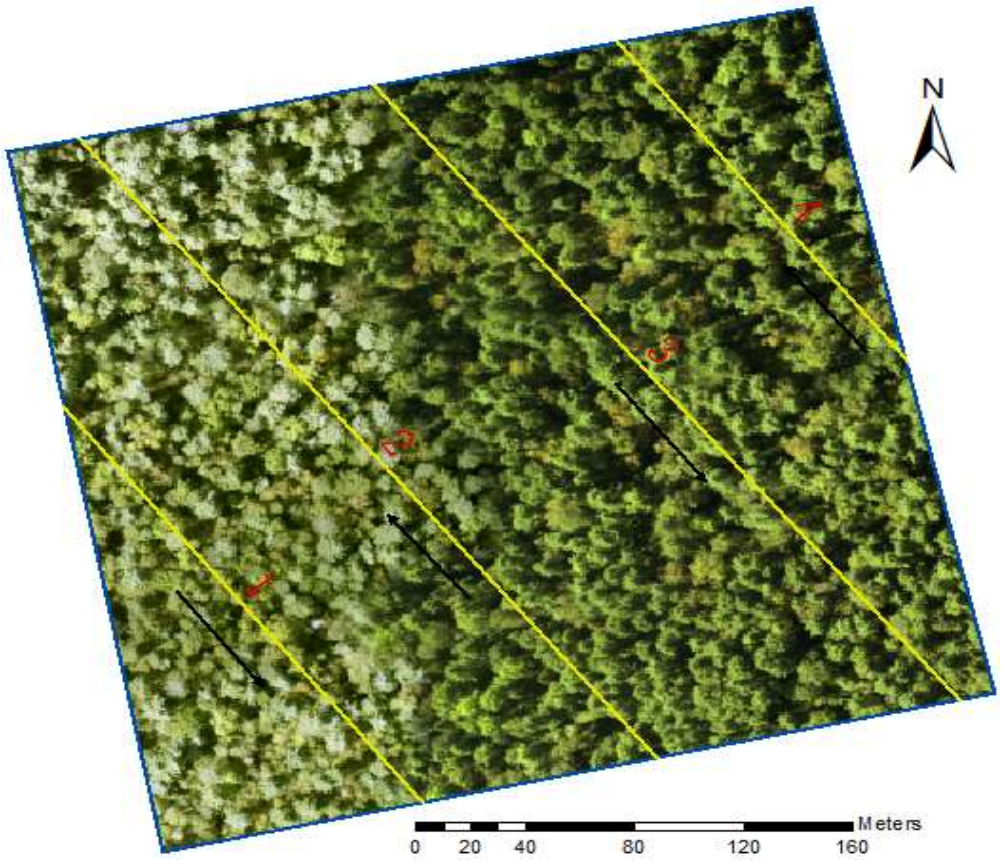


Figure 1 The CCD image of test area and centerlines of the three flight strips

2.2 LIDAR data

The LIDAR data acquisition was performed in the summer of 2009 at a leaf-on condition utilizing the Riegl LMS-Q560 scanner, a small footprint commercial LIDAR system with full waveform digitizer. The waveform digitizer samples at a rate of 1 ns time intervals at 1,550 nm and records both transmitted and backscattered pulses. The platform was a Chinese YUN-5 aircraft, flying at an altitude of 1,000 m, about 700 m high. The scan angle was ±22.5° for a ground swath of about 700 m. Laser point spacing was about 1 m, and the laser beam divergence was 0.5 mrad. Ten flight lines with about 80% overlap were designed to cover 1,600 ha of the center of the nature reserve. The data collection configuration resulted in a point density of about three points m-2. In addition, a 3-band (Red, Green, and Blue) high-resolution CCD camera was used to capture aerial image at the same time of the LiDAR scanning mission. So the dataset contains a geometrically corrected orthophoto with 0.2 m spatial resolution, for the assistant accuracy assessment and training data selection of tree species classification.

The plot is covered by four adjacent flights LIDAR data and black arrows show the flying direction of the plane in the Fig.1. The strip 2 and 3 cover the entire plot and the strip 1 and 4 cover almost 80% of the plot. The yellow lines are the centerlines of the four flights.

3. METHODOLOGY

Our concept for tree species classification comprises the following steps. First, the waveforms are decomposed into a series Gaussian pulse to obtain 3-dimensional points and their attributes. Then a relative calibration method is applied to calibrate the amplitude and energy. Second, individual tree crowns are delineated by a watershed segmentation of the CHM with a crown control method. The next step is to extract features including the number of reflections per waveform, the calibrated amplitude and energy, the pulse width of each tree. Then tree species are classified using the extracted features with a SVM classifier. Finally, the results are verified with experiment truth data and the impact of the relative calibration of the LIDAR data on tree species classification is evaluated.

3.1 Gaussian decomposition

The objective of waveform processing is to extract more information from raw data than provided by the system. In general, small-footprint waveform can be modeled reasonably by a series of Gaussian pulses (Wagner, W. et al., 2006). Gaussian decomposition is the most widely applied processing approach currently. One return waveform can be decomposed into its Gaussian components in this form:

$$f(x) = b + \sum_{i=1}^{n} a_i e^{-(x-x_i)^2/2\sigma_i^2} \quad (1)$$

where $f(x)$ represent the received waveform, b is the noise level of the waveform, n is the number of Gaussian components, a_i, x_i and σ_i denote the amplitude, position and half-width (standard deviation) of each Gaussian respectively. In this paper, a non-linear least-squares method with the Levenberg-Marquardt algorithm was used to fit the return waveforms.

3.2 Relative calibration

Although the received waveform is affected by many factors, some studies have shown that waveform energy and amplitude decrease with an increase of the range between the sensor and target, which is considered to be the most important influence factor of waveform data. In this study, the LIDAR system of Riegl has a peculiar function of recording the transmitted pulses. Some differences were found by comparing the transmitted pulses. The differences were mainly reflected in changes of energy, and the pulse widths were almost the same. Thus, the range and transmitted pulse changes were considered during relative calibration and other factors were neglected because of their difficulty to obtain.

Transmitted pulse energy variation would inevitably have an impact on back echoes energy; to restrain this influence, a standard pulse energy needs to be set to normalize all of the transmitting pulse. Thus, the normalization factor can be described as:

$$C_{Index} = \frac{P_{t\_sta}}{P_t} \tag{2}$$

where C_{Index} is the normalization factor, p_{t_sta} is the standard transmitted pulse energy and p_t is the transmitted pulse energy. In this study, p_{t_sta} was set as the average of all of the transmitted pulses energy.

According to the LIDAR equation and the different scattering properties of objects, the relative calibration formula based on range for each return energy of Gaussian decomposition result can be described as (refer to Guangcai Xu et al. (2011) for more detail regarding the formula derivation):

$$p_{cal} = p_r * C_{Index} * (\frac{R}{R_{cal}})^n \tag{3}$$

where p_{cal} is the calibrated energy, p_r is the energy of decomposed Gaussian pulse, C_{Index} is the transmitted pulse normalization factor, R is the range of sensor to object, R_{cal} is the reference range and n is the calibration factor with values of 2, 3 or 4.

The correction factor is determined by the scattering properties of the target. In the correction process, the value of n was varied from 2 to 4 at a step of 0.01, and the correlation coefficient of the calibrated values and corresponding real range was calculated. Then the value of n at the minimum correlation coefficient was selected as the calibration factor. R_{cal} was set as the same value of the mean range of all flights.

Following the determination of n and the combination of the Gaussian function expression, each calibrated amplitude of Gaussian decomposition result can be computed as:

$$a_{i\_cal} = a_i * C_{Index} * (\frac{R}{R_{cal}})^n \quad (4)$$

where a_{i_cal} is the calibrated amplitude and a_i is the amplitude of the Gaussian decomposition result.

3.3 Individual Tree Isolation from LiDAR Data

In previous research, a morphological crown control based watershed algorithm was proposed to isolate individual trees. The algorithm could gain individual tree positions, heights and crown radius. The crown control, which determines the real tree crown area approximately, used to limit the watershed operation locating in crown area. This could restrain the over segmentation, which was commonly encountered in most individual tree isolation algorithms. In the algorithm, the morphological crown control was imported to determine outer markers. The local maxima method with outer markers is used to identify potential tree positions. Double watershed transformations, in which a reconstruction operation is inserted into these two transformations, are then applied to delineate the tree crown. Finally, the individual trees are isolated, and their parameters are extracted.

Outer markers were used to eliminate non-crown parts from the CHM. The crown control applied a morphological closing operator, which smoothes the object outliners, fills the narrow gaps as well as long and slender chasms, removes small holes and fills in outliner breaks to the CHM using a near circular structural element. The inner markers were used to mark potential tree positions in the CHM. The local maximum was frequently used to identify potential treetops in the LiDAR-derived CHM. But, this resulted in under-recognition. Here, local maxima detection with outer markers was used. Local maxima meant the external boundary pixels of object pixel all had a lower value. As a result, potential treetops could be found properly.

The 1st watershed transform was applied to the smoothed and negated CHM with both inner and outer markers.

$$CHM_{Watershed} = Mask(WaterShed(CHM, inMark), outMark) \quad (5)$$

where $WaterShed()$ is the watershed function, $Mask()$ is the mask function, CHM is the smoothed and negated CHM, $inMark$ is the inner markers, $outMark$ is the outer markers, and $CHM_{Watershed}$ is the 1st watershed transform result.

The 1st watershed transform results had many over-segmentation problems; therefore, a reconstruction was applied to reconstruct the smoothed CHM. Here, the reconstruction filled the basins without inner markers with values that were determined by the corresponding basin boundaries. For the reconstruction, every basin in the $CHM_{Watershed}$ was checked for whether it had inner markers. If the basin did not have inner markers, then the basin pixels in the CHM were replaced by the mean value of the watershed line pixels of the corresponding basin in the CHM.

$$CHM(Index_{Basins\_Without\_inMark}) = Mean(CHM(Boundary\_Index_{Basins\_Without\_inMark})) \quad (6)$$

where $Index_{Basins\_Without\_inMark}$ is the index of the basin pixels without an inner marker inside, $Mean()$ is the mean value function, and $Boundary\_Index_{Basins\_Without\_inMark}$ is the index of the watershed line pixels of the corresponding basin.

The 2nd watershed transform was applied to the reconstructed CHM with the outer and inner markers. After the 2nd watershed transform, the results were much better than the 1st watershed transform results. Most of the basins had inner markers inside, and each basin represented the corresponding crown area of an individual tree. Finally, the individual tree parameters could get from these basins combined with point cloud data.

3.4 Classification

Basing on the relative calibration and the parameters obtained from Gaussian decomposition and tree crowns bound delineated by a segmentation of the CHM, characteristic features of the individual trees are extracted. All the points with the attributes, such as the energy, amplitude, calibrated energy and amplitude, pulse width, the number of return waveforms, in the range of a tree were collected for future classification. Some statistics, such as the maximum, mean and minimum values, were calculated as the feature of a single tree for species classification. However, not all of the statistics satisfied the goal of species discrimination and some ill-suited statistics may reduce the separability of different tree species. The most important variables were selected in the pre-classification process, which is listed in table 1. In the classification process, the standard deviation, amplitude, energy and the number of all return echoes are combined as a group for future species classification. Finally, two combination groups (1, 2, 3, 6 and 1, 4, 5, 6) were employed to distinguish various trees.

Table1. Extracted waveform LIDAR parameters of single trees

Attribute	Waveform feature of all returns
1	Average standard deviation
2	Average amplitude
3	Average energy
4	Average Calibrated amplitude
5	Average Calibrated energy
6	The number of return echoes

Since overtopped and suppressed trees can normally not be identified on CHM images, only the first two crown classes, being dominant and co-dominant were considered for further training and testing data collection. So the main tree species of the first two dominant trees (Pine, Spruce and Fir, Elm, Birch, Linden and Chinese ASH) are selected as the object of classification.

Training data selection is of great importance to the classification. Some field data that can perfectly match the result segmentation were selected as ROIs (region of interests). Then the ROIs was divided into two sets including a training set for classification and a ground truth set for classification result evaluation. At last a SVM classifier was employed to distinguish various trees.

4. Results

4.1Gaussian decomposition and relative calibration

The Gaussian decomposition method was applied to the four flight tracks over the study area. In total, 609567 waveform profiles were processed and 904397 points were extracted with an average point density approximately 10 point m-2. About 99.89% of the data were successfully fitted by Gaussian pulses; of those results, 67.40% were first echoes, 25.75% were second echoes, 5.72% were second echoes, and 1.13% showed more than three echoes. The maximum number of return is eight and 38.29% of the pulses can get more than one return points. This is mainly due to the complex vertical structure in forestry area. The number of points and return echo numbers derived from the Gaussian decomposition was greater compared to the real-time processing echo extraction by the system. Some extra parameters such as the position, the number of echoes and the sequence of return echoes within the beam, amplitude, standard deviation, energy and range between the sensor and object of each decomposed Gaussian waveform.

Basing on the Gaussian decomposition, the mean value of range and transmitted pulse energy of the four flights were calculated. Then formula 3 and 4 were adopted to calibrate the amplitude and energy, and the correction factors n at the minimum correlation coefficients of the four flight lines were 2.01, 2.03, 1.97 and 2.08, which were close to the properties of the extended targets.

4.2 Individual Tree Isolation

An invalid values filling approach with morphological crown control, as in one of our previous studies (Dan Zhao et al., 2012), was applied. After invalid value filling, individual tree isolation method was applied, and the results are shown in Figure 2. By visual judgment, most of the canopy crowns were assigned to the specified individual trees. The gray crosses show the isolated individual tree positions. A total of 2188 trees were found and marked in the large CHM, and 1847 trees were located in the sample plot. From the figure, the most isolated individual trees had a relatively smooth crown shape, which was close to the real tree crown shape. The tree height distribution of the 1847 trees is shown in Figure 3. From the figure, more than 90% of the trees found by the algorithm had a tree height greater than 15 m.

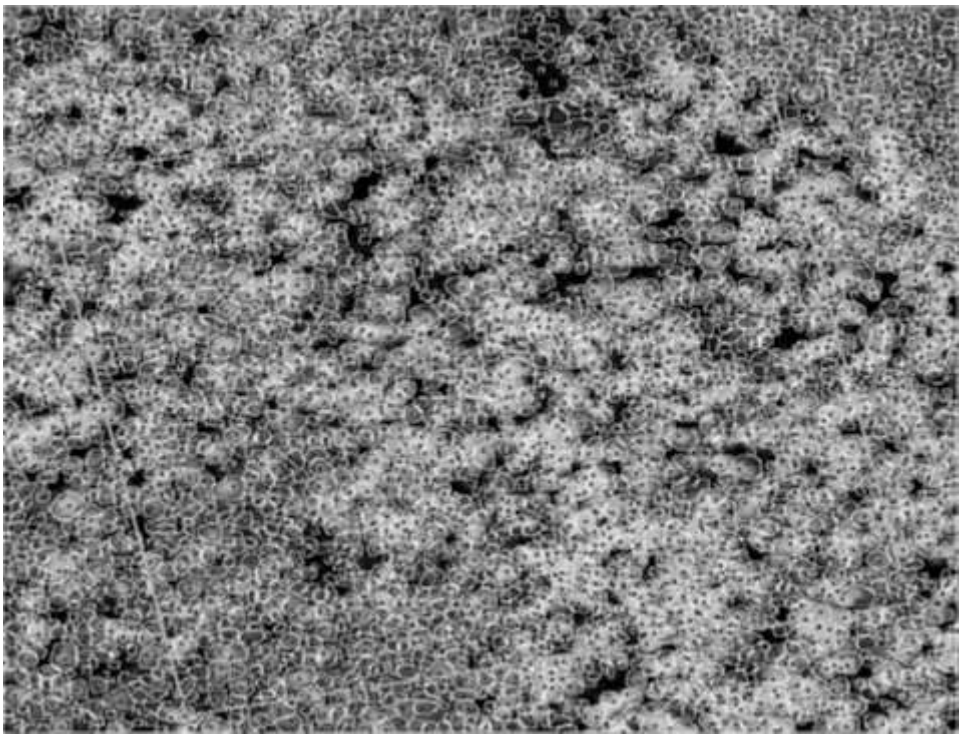


Figure 2 The individual tree isolation result of the forest sample plot

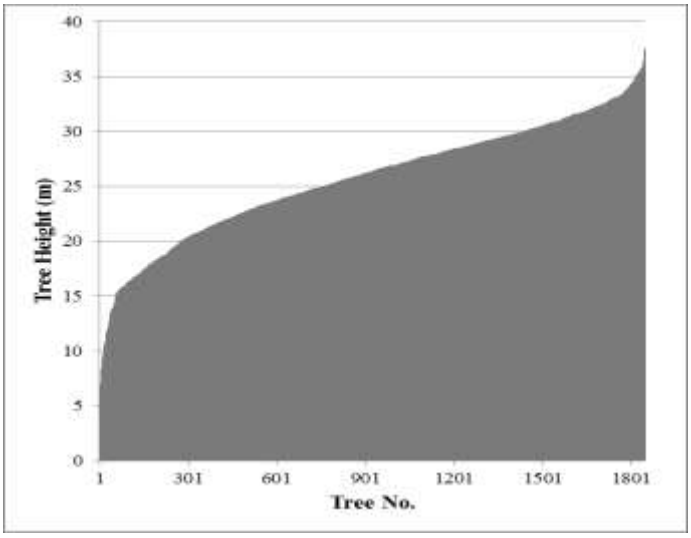


Figure 3 Tree height distribution of 1847 isolated trees in the forest sample plot

4.3 Classification

In the classification process, two combination groups (the uncalibrated data and calibrated data) were employed to distinguish various targets with a SVM classifier. Three classification depths (seven species, five species and two species) were applied to evaluate the capability of parameters derived from waveform data in tree species discrimination and to show the impact of species identification after calibration. The overall classification accuracy, kappa coefficient, producer's accuracy of classification results are listed in table 3. The total accuracy of three classification depths was improved with the tree species reducing. Following a comparison of the two classification groups, the classification accuracy of the relative calibrated data increased steadily when compared to the

uncorrected data.

Table 2: Classification accuracy of uncalibrated and calibrated data of three classification depths and the value are given in percent. RP= Red Pine, ML =Manchurian linden, WB=White birch, KS=Koyama Spruce, MA= Manchurian ash, Con= coniferous, Dec=deciduous.

Uncalibrated data

Ground truth							
	RP	**ML**	**Fir**	**WB**	**KS**	**Elm**	**MA**
RP	73.18	0	18.22	22.26	7.02	17.01	45.32
ML	0	50.8	14.18	7.79	1.21	0	3.52
Fir	0	15.19	64.72	9.27	22.54	16.45	0
WB	5.92	0	0	44.58	12.77	3.47	25.3
KS	8.17	22.09	2.89	0	27.6	15.44	6.86
Elm	12.73	7.85	0	13.35	23.12	47.62	0
MA	0	4.07	0	2.74	5.75	0	19
Overall accuracy= 49.97% Kappa coefficient= 0.4029							

Ground truth						
	RP	**ML**	**Fir**	**WB**	**Elm**
RP	62.92	0	3.75	0	30.37
ML	0	62.14	10.76	7.79	0
Fir	25.17	22.38	68.99	9.27	42.96
WB	7.25	11.92	16.5	69.58	6.19
Elm	4.66	3.56	0	13.35	20.48
Overall accuracy= 60.40% Kappa coefficient= 0.4872					

Ground truth			
	Con	**Dec**
Con	84.1	22.09
Dec	15.9	77.91
Overall accuracy= 82.16% Kappa coefficient= 0.5995		

Calibrated data

Ground truth							
	RP	**ML**	**Fir**	**WB**	**KS**	**Elm**	**MA**
RP	70.57	0	21.1	20.25	31.74	11.47	5.38
ML	4.67	69.33	14.18	7.79	1.21	6.1	0
Fir	0	15.19	47.21	9.27	23.17	10.35	3.52
WB	0	0	0	41.69	0	0	13.44
KS	16.7	0	17.51	0	32.55	17.46	6.86
Elm	5.45	11.41	0	0	7.3	51.15	0
MA	2.61	4.07	0	20.99	4.03	3.47	70.81
Overall accuracy= 55.07% Kappa coefficient= 0.4029							

Ground truth					
	RP	**ML**	**Fir**	**WB**	**Elm**
RP	75.24	0	6.03	2.39	17.01
ML	0	62.14	6.7	5.89	0
Fir	15.45	22.38	71.01	16.19	34.97
WB	3.85	11.92	11.76	68.11	3.47
Elm	5.45	3.56	4.49	7.42	44.54
Overall accuracy= 66.15% Kappa coefficient= 0.5631					

Ground truth		
	Con	**Dec**
Con	81.2	10.02
Dec	18.8	89.98
Overall accuracy=85.72% Kappa coefficient= 0.7134		

5. DISCUSSION & Conclusion

5.1 Discussion

Many effective methods of LIDAR data classification and calibration have been developed by previous research studies. Nevertheless, very few authors have investigated the impact of radiometric calibration of LIDAR data on single tree species recognition. A complete workflow including Gaussian decomposition, relative calibration, individual tree isolation and tree species classification was described. The individual tree positions and tree crowns were acquired using a morphological crown control based watershed algorithm. A simple relative radiometric calibrated method based on

range and normalization of transmitted pulses was introduced to normalize the Gaussian decomposition results of full-waveform data. The paper addresses the impact of the relative radiometric correction on the accuracy of single tree species classification and the result shows that the relative calibration method could improve the capability of tree species identification.

This paper proposed a morphological crown control based watershed method to isolate individual trees and to obtain their parameters. The morphological crown control was an approach that only processed the individual tree isolation in the area of crown instead of the whole CHM. This approach could improve the accuracy of the results and the algorithm efficiency. Our algorithm produced a good result in our experiments. However, some problems were encountered, and some improvements were required. For the determination of the outer markers, some other methods were tried, such as the Gaussian Laplacian Operator and the difference of Gaussian, but the results were not as good as the morphological crown control. Another proposal was that corresponding optical images that were acquired with the LiDAR data could provide crown areas more precisely. In this step, the threshold $threshold_{Crown\_Control}$ was the key point. How to determine the threshold in different forest types and different forest status requires additional study.

In this study, relative radiometric calibration of full-waveform small-footprint LIDAR was performed based on a radar equation, but absolute calibration was not carried out due to the lack of calibrated reference targets. Although the calibration approach achieved relatively effective results for object classification, there are still some problems worthy of further study. The correction formula was found on the assumption that the surface is a Lambertian reflector. Because most objects in nature are not ideal, Lambertian and its radiation are always related to the angle of incidence. The enhancement of the comparability of multiple echoes and single echo requires in-depth mechanistic research of LIDAR. In short, the current waveform data correction theory is not perfect, and a unified standard has not been developed. The airborne LIDAR waveform data correction model depends on LIDAR mechanistic studies, and correction theory development is also conducive to the improvement of the mechanism model. The motivation behind this research was to introduce the incidence angle and the number of returns into the calibration formula to improve the correction result of waveform data

5.2 Conclusion

This study focused on the airborne full-waveform LIDAR data processing, relative calibration and single tree species discrimination with parameters extracted from waveform. The overall results showed that the parameters acquired from waveform could distinguish different tree species. The following important aspects can be concluded:

- The waveform Gaussian decomposition increases the spatial point density and provides additional attributes for the detected points.
- The results show that individual tree isolation algorithm can isolate most dominant and subdominant trees (tree heights larger than 15 m in the plots), more than half of the intermediate trees (tree heights between 10 and 15 m in the plots) and some suppressed trees (tree heights less than 10 in the plots).
- The three parameters (amplitude, echo width and energy) extracted from small-footprint full-waveform of Gaussian decomposition can provide valuable information for tree species discrimination.

- Calibration of full-waveform data is necessary for its application in object classification. The parameters of the groups of the corrected data improved classification accuracies by an additional 3.56% to 5.75% compared to the uncalibrated data for different classification depths.

6. Acknowledgements

This study was supported by Chinese National 863 Research Program 'The key technologies of remote sensing estimation in the global forest biomass and carbon stocks' (2012AA12A306).

7. REFERENCES

Briese Christian, Bernhard Höfle., Hubert Lehner., Wolfgang Wagner., Martin Pfennigbauer and Andreas Ullrich. Calibration of full-waveform airborne laser scanning data for object classification. *Proc*. *SPIE 2008*, 6950, 69500H.

Dan Zhao, Yong Pang, Zengyuan Li, Guoqing Sun. Filling Invalid Values in a LiDAR-Derived Canopy Height Model with Morphological Crown Control. *International Journal of Remote Sensing*. 2011. (Acceptance)

Doneus, M. and Briese, C., Digital terrain modeling for archaeological interpretation within forested areas using full-waveform lasers canning, *The 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST* , 2006,155 - 162 .

Hofle, B., Hollaus, M., Lehner, H., Pfeifer, N. and Wagner, W., Area-based parameterization of forest structure using full-waveform airborne laser scanning data. *Proceedings of the Silvilaser*, 2008, 2, 227-235.

Hollaus, M., Mücke, W., Hofle, B., Dorigo, W., Pfeifer, N., Wagner, W., Bauerhansl, C., Regner, B. Tree species classification based on full-waveform airborne laser scanning data. *Proceedings of Silvilaser*, 2009, 54-62.

Johannes Heinzel and Barbara Koch. Exploring full-waveform LiDAR parameters for tree species classification. *International Journal of Applied Earth Observation and Geoinformation*, 2011, 13(1), 152-160.

Mallet, C., Bretar, F., 2009. Full-waveform topographic LIDAR: state-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing* 64, 1-16.

Neuenschwander Amy L., Lori A. Magruder and Marcus Tyler. Landcover classification of small-footprint, full-waveform lidar data. *Journal of Applied Remote Sensing*, 2009, 3, 033544.

Persson, A., Söderman, U., Töpel, J. and Ahlhberg, S., Visualisation and analysis of full-waveform airborne laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 3/W19), 2005, 103-108.

Reitberger, J., Krzystek, P. and Stilla, U. Combined tree segmentation and stem detection using full waveform LIDAR data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part3/W52), 2007, 332-337.

Reitberger, J., Krzystek, P. and Stilla, U. Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing*, 2008, 29 (5), 1407-1431.

Ullrich, A., Hollaus, M., Briese, C., Wagner, W. and Doneus, M.Utilization of full-waveform data in airborne laser scanning applications. *Proc*. *SPIE 6550*, 2007, 65500S.

Wagner, W., Hollaus, M., Briese, C., Ducic, V. 3D vegetation mapping using small-footprint full-waveform airborne laser scanners. *International Journal of Remote Sensing*,2008, 29, 1433‐1452.

Non-Parametric Point Classification for Phase-Shift Laser Scanning

Glenn Newnham[1], John Mashford[1], Pyare Püschel[2], John Armston[3,4], Darius Culvenor[1], Anders Siggins[1], Mattias Nyström[5], Nicholas Goodwin[3] & Jasmine Muir[3]

[1] CSIRO Land & Water and Sustainable Agriculture Flagship, glenn.newnham@csiro.au, john.mashford@csiro.au, darius.culvenor@csiro.au,
[2] University of Trier, p.pueschel@uni-trier.de
[3] Remote Sensing Centre, Queensland Department of Science, Information Technology, Innovation and the Arts, John.Armston@derm.qld.gov.au, Nicholas.Goodwin@derm.qld.gov.au, Jasmine.Muir@derm.qld.gov.au
[4] Joint Remote Sensing Research Program, School of Geography, Planning & Environmental Management, University of Queensland
[5] Swedish University of Agricultural Sciences, Mattias.Nystrom@slu.se

Paper Number: SL2012-090

Abstract

The classification of data points in a laser scan recorded in a forest environment may help to distinguish object types such as the ground surface, woody components and foliage, from which terrain models, stem characteristics and foliage profiles can be derived. Phase-shift laser scanner data presents a particularly challenging problem in forest environments because of inherently noisy return ranges and the lack of unambiguous response to non-interceptions of the beam. In this case the basic challenge is to distinguish true gaps in the canopy from those returns emanating from solid objects such as vegetation components and the surrounding terrain surface. Here we present a non-parametric framework for classifying phase-shift laser scanner data into gaps and solid object types relevant for the assessment of vegetation structure. The method is applied to a laser scan recorded in a Eucalyptus forest in Brisbane Australia. We show that the differentiation of grass, woody material, foliage, sky and sun sensor artefacts can be achieved, although some confusion between object classes is apparent.

1. Introduction

Phase-shift laser scanners employ a constant beam with intensity modulated at a series of frequencies. The shifts in phase of any returned signal for each frequency are used to determine the range of the target. The main benefit of phase-shift scanners is that they can determine range more rapidly than time-of-flight instruments; in the order of 1 million points per second. Current systems are also generally lighter and cheaper than "equivalent" time-of-flight instruments. This comes at the cost of effective range, which is generally limited to less than 100m depending on the laser pulse peak energy, wavelength and target reflectance characteristics.

Phase-shift scanners have been applied to measurement of canopy structure (Antonarakis, 2011; Balduzzi et al., 2011; Park et al., 2010) and the assessment of woody vegetation components (Bienert et al., 2007; Liang et al., 2008; Park et al., 2010). However, the phase-shift ranging method is subject to ambiguity where a single outgoing beam is intercepted by multiple targets or where the beam is not intercepted within a range limit, dictated by signal to noise ratio. Often a laser pulse into a clear beam path will result in random ranges being returned. This is generally addressed in firmware or through post-processing, with points removed based on the

smoothness, or surface roughness with respect to neighbouring laser pulses.

In traditional applications of TLS for built environments such as engineering and architecture, multiple interceptions and non-interceptions can be dealt with by removing points from the point cloud. In vegetation structural analysis, these two cases need to be dealt with differently, since the removal of points that indicate multiple hits would overly inflate gap probability estimates and consequently decrease apparent stem size, foliage cover and leaf area index (LAI). Sky points on the other hand need to be removed reliably so that true gaps can be identified.

Phase-shift scanner systems generally include a number of filters to deal with spurious points which are implemented in firmware or post-processing software. These include simple threshold filters for range and intensity and textural filters that specify the minimum proportion of neighbouring points that must fall within a given distance from each point. However, these filters are designed to remove both points resulting from multiple interceptions and points resulting from non-interceptions. In this paper we demonstrate a non-parametric point classification method used to discriminate between laser returns from complex vegetation canopy media and non-interceptions within canopy gaps.

2. Methods

The data used in this study were recorded in a forest plot (100 m x 50 m) in the D'Aguilar National Park, near Brisbane Australia on the 8th of November 2011. The site is a Eucalyptus woodland with very little understorey and a ground layer of grass and woody debris. The tree heights within the plot have a mean height and diameter at breast height (DBH) of 13.8 m and 0.18 m respectively. The maximum height within the plot was 37 m and the maximum DBH was 1.05 m.

A laser scan was recorded using a FARO Focus 3D 120 phase-shift scanner. The FARO uses a 905 nm laser with beam divergence of 0.16 mrad, which scans at a rate of 976000 pulses per second. The nominal maximum range for the instrument is 120 m. In addition to range, the instrument records the laser return intensity, and measures of red, green and blue passive reflectance using an integrated camera system. In addition to these data, a measure of surface roughness t was computed from the raw range information r by taking the natural log of the sum of differences in range to adjacent laser returns as follows:

$$t_{i,j}(r) = \ln\left[\sum_{d\theta=-1}^{1}\sum_{d\varphi=-1}^{1} | r_{i,j} - r_{i+d\theta,j+d\varphi} |\right] \quad (1)$$

where $d\theta$ and $d\varphi$ indicate scan increments in the zenith and azimuth directions respectively. Let the data associated with each return in the resulting point cloud be denoted by the vector y = (range, intensity, surface roughness) and the discrete variable u range over the possibilities of the objects encountered in the direction of the beam, so $u \in X$ where X = {Vegetation, Canopy Gaps, ... }. Using Bayes' theorem, the probability of the object being of type u given the data y is given by:

$$p(u \mid y) = \frac{p(u)p(y \mid u)}{p(y)} \tag{2}$$

where

$$p(y) = \sum_{u \in X} p(u)p(y \mid u) \tag{3}$$

In this case $p(u)$ is the prior probability of the return being from target type u and $p(y \mid u)$ is the conditional probability of the value y given the object being of type u. We can estimate the probability of the object being of type u given n input variables by applying Bayes' theorem sequentially. We first estimate prior probabilities and then generate posterior probabilities based on the information from the first point data variable y_1 (e.g. range). These posterior probabilities can then be used as prior probabilities for a subsequent estimate based on the next data variable (e.g. intensity). Let $p_1(u \mid y_1)$ be the posterior probability of the point being of type u given the first variable y_1. Then

$$p(u \mid y_1) = \frac{p(u)p(y_1 \mid u)}{p(y_1)} \tag{4}$$

Now we use $p(u \mid y_1)$ as the prior probability to estimate $p(u \mid y_1, y_2)$ based on the data distribution of the second variable y_2 as follows.

$$p(u \mid y_1, y_2) = \frac{p(u \mid y_1)p(y_2 \mid u)}{p(y_1, y_2)} = \frac{p(u)p(y_1 \mid u)p(y_2 \mid u)}{p(y_1, y_2)} \tag{5}$$

Extended to the case of n variables, the estimate of the probability of the point being of type u can be calculated by taking the product of the data distribution probabilities as follows

$$p(u \mid y) = p(u \mid y_1, y_2, \ldots, y_n) = \frac{p(u)\prod_{i=1}^{n} p(y_n \mid u)}{p(y)} \tag{6}$$

where

$$p(y) = \sum_{u \in X} p(u)\prod_{i=1}^{n} p(y_i \mid u) \tag{7}$$

In this study, the objects u which form the set X were based on the major elements of the forest plot and included grass, woody material, foliage, canopy gaps and the sun sensor. Bare ground was not included in X due to the high level of grass cover. The sun sensor was included because the instrument has a detector protective mechanism where ambient light exceeds a certain threshold. This causes unexpected range and intensity data when sampling in the direction of the sun. This region also needs to be treated differently to canopy gaps in the analysis of vegetation

structure as effectively there is no measurement within this region and it needs to be ignored in the calculation of structural metrics such as LAI.

Prior probabilities for each u were considered equivalent for each object type such that:

$$p(u) = \frac{1}{\#X} = \frac{1}{5} \tag{8}$$

Final classification was based on the maximum probability over all possible object types (u). The accuracy of the classification was assessed in terms of individual accuracies for each class, the total percentage of the training set correctly classified and the kappa statistic (Cohen, 1960), which is a more conservative estimate of accuracy taking into consideration agreement that can occur by chance.

3. Results

The laser scan was recorded at an angular sample spacing of 3 mrad resulting in approximately 1.7 million pulses across the full hemispherical scan. Estimates of $p(y \mid u)$ were based on subjective selection of training regions within a Plate Carrée projection of the laser scan (see Figure 2). A total of 3.5% of the scan was selected as training data, divided into each of the object categories as detailed in Table 1.

Table 1: Number of laser pulses used in each training set

Class	Grass	Woody	Foliage	Sky	Sun Sensor
Pixels	20497	20687	9566	7794	4268
Percentage	1.2	1.2	0.5	0.4	0.2

Marginal probabilities for range (Figure 1a) highlight the difficulty selecting angular regions of vegetation components at ranges greater than 25 m. Grass was generally selected within 10 m from the instrument, woody material was generally selected within 10 m and foliage in the tree crowns within 25 m. The canopy gaps and solar disk training regions showed some local peaks but were generally more evenly distributed out to the maximum recorded range of 153 m. The non-normal characteristics of these data provide good examples of why parametric methods struggle to adequately characterise these distributions.

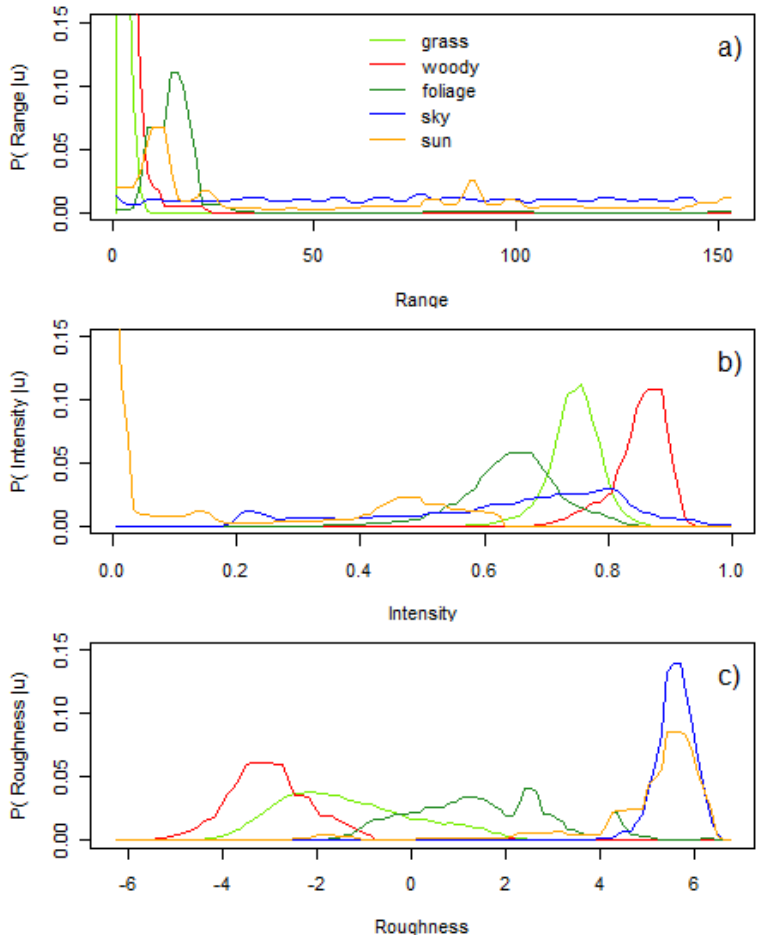


Figure 1: Marginal probabilities $p(y \mid u)$ for each of the object types recorded in the laser scan, where u is equal to a) range from the scanner, b) laser return intensity and c) a surface roughness parameter.

The intensity of the laser return pulse is dependent on the range to the intercepting object, its reflectance at 905 nm and the proportion of the beam intercepted. The marginal probabilities for intensity (Figure 1b) show a clear distinction between the solar disk and other objects within the scan. However, the intensity for canopy gaps (sky) is more difficult to discriminate from both woody and non-woody vegetation components. Of the vegetation components, woody vegetation showed the highest intensities, while crown foliage showed the lowest intensities.

Surface roughness showed clear distinctions between vegetation components and the canopy gaps and solar disk (Figure 1c). The near random ranges for these non-vegetation classes produced high measures of roughness, even when a logarithmic transform is applied. The lowest measures of roughness were associated with the woody vegetation components due to generally smooth surfaces on trunks and branches.

The application of Eq. 6 over the complete laser scan produced a realistic classification of the data (Figure 2). Confusion between classes can be seen at the edge of woody material, which are often classified as grass. This is highlighted by an 8% misclassification in the confusion matrix shown in Table 2. There are also laser returns within the ground cover layer classified as woody material (6% in Table 2), although it is possible some of these may be associated with woody debris and bare ground. In the far field both trunks and grass often appear to be classified as foliage. Some misclassification of sky as the sun sensor class is also apparent and supported by a 3% misclassification shown in Table 2.

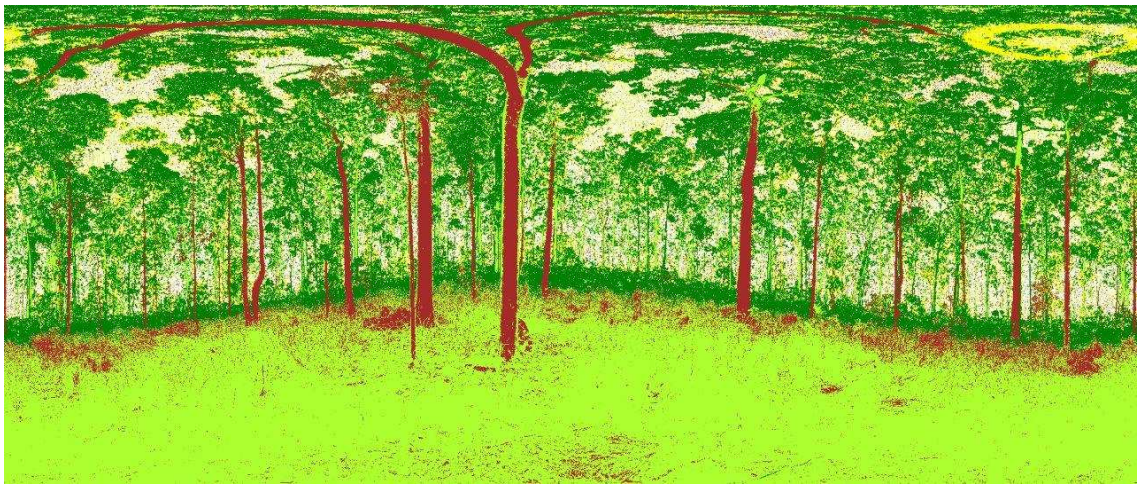


Figure 2: Phase-shift laser scanner classified as grass (light green), woody material (brown), foliage (dark green), sky (white) and solar disk (yellow).

Overall percentage of the training data for all classes classified correctly was 94% and the kappa statistic was 0.78. This compares favourably to a simple Gaussian minimum distance classifier which produced an accuracy of 91% and a kappa statistic of 0.65.

Table 2: Confusion matrix for the non-parametric classifier

Class	Grass	Woody	Foliage	Sky	Sun Sensor
Grass	94	8	0	1	0
Woody	6	90	0	1	0
Foliage	0	2	99	1	3
Sky	0	0	0	94	19
Sun Sensor	0	0	0	3	78
Total	100	100	100	100	100

4. Discussion and Conclusions

The advantage of the Bayesian classification method described in the paper is that is can deal with any distribution of input data. This has advantages in the case of phase-shift laser scanning data, where non-interceptions of the laser can result in unusual distributions of both range and intensity. The method described also has the advantage that it can be used to produce posterior probabilities for input parameters which may be used in the classification of further scans recorded using the same scanner type. The method may be extended to include a second iteration in the calculation of the posterior probability which includes a prior-probability for spatial context. In the simplest case this probability could be based on the probabilities assigned to adjacent pixels, but might also include more complex morphological filters that consider known structural forms such as the near form of trunks and the connectivity of branches.

There are a number of limitations of the current implementation of the methods. Input data used in this study was not necessarily ideal given our knowledge of the objects being classified. For example, the intensity expected from a given vegetation component type (woody, non-woody) is

a function of range and it may be useful to incorporate an input variable such as apparent reflectance (Jupp et al., 2008) as a substitute. Given the objects selected in this study, an input variable that describes height above the ground surface may also be valuable.

The method described uses the product of marginal probabilities for estimation of the posterior probability for a given object class. A further refinement of the methods would be to construct a full joint probability distribution for variables known prior to the first iteration of the Bayesian computation. This is likely to account for interdependences between variables more effectively. Further developments will also need to include validation of the method using independent data.

The objective of this study was the filtering of phase-shift laser scanning data, where it is often difficult to distinguish between vegetation and gaps using traditional rule based approaches. However, as demonstrated, the method has broader application in classification of laser scanner point clouds, such as the discrimination of woody and non-woody vegetation components. This is generally being tackled using deterministic methods (Bienert et al., 2007; Liang et al., 2008; Park et al., 2010). However, incorporating these deterministic methods as input to a probabilistic framework, such as that described here, may be a powerful solution that not only allows discrimination of vegetation component types within laser scanner data, but also assigns confidence to the classified data.

References

Antonarakis, A.S. (2011). Evaluating forest biometrics obtained from ground lidar in complex riparian forests. Remote Sensing Letters, 2, 61-70

Balduzzi, M.A.F., Van der Zande, D., Stuckens, J., Verstraeten, W.W., & Coppin, P. (2011). The Properties of Terrestrial Laser System Intensity for Measuring Leaf Geometries: A Case Study with Conference Pear Trees (Pyrus Communis). *Sensors*, 11, 1657-1681

Bienert, A., Scheller, S., Keane, E., Mohan, F., & Nugent, C. (2007). Tree detection and diameter estimations by analysis of forest terrestrial laserscanner point clouds. In, *ISPRS Workshop on Laser Scanning* (pp. 50-55). Espoo, Finland

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20 (1): 37-46.

Jupp, D.L.B. et al., 2008. Estimating forest LAI profiles and structural parameters using a ground-based laser called Echidna. *Tree Physiology*, 29(2): 171-181.

Liang, X., Litkey, P., Hyyppa, J., Kukko, A., Kaartinen, H., & Holopainen, M. (2008). Plot-level trunk detection and reconstruction using one-scan-mode terrestrial laser scanning data. 2008 *International Workshop on Earth Observation and Remote Sensing Applications*, 136-140

Park, H., Lim, S., Trinder, J., & Turner, R. (2010). 3D Surface Reconstruction of Terrestrial Laser Scanner Data for Forestry. *IEEE International Geoscience and Remote Sensing Symposium*, 4366-4369

Comparing Time-of-Flight and Phase-Shift Terrestrial Laser Scanners for Characterising Topography and Vegetation Density in a Forest Environment

Glenn Newnham[1], Nicholas Goodwin[2], John Armston[2,3], Jasmine Muir [2] & Darius Culvenor[1]

[1] CSIRO Land & Water and Sustainable Agriculture Flagship, Glenn.Newnham@csiro.au, Darius.Culvenor@csiro.au
[2] Remote Sensing Centre, Queensland Department of Science, Information Technology, Innovation and the Arts, John.Armston@derm.qld.gov.au, Nicholas.Goodwin@derm.qld.gov.au, Jasmine.Muir@derm.qld.gov.au
[3] Joint Remote Sensing Research Program, School of Geography, Planning & Environmental Management, University of Queensland

Paper Number: SL2012-095

Abstract

Time-of-flight terrestrial laser scanners are often designed specifically for long range outdoor use, while phase-shift scanner design has generally focussed on high speed indoor and short range use. Applications of laser scanners for vegetation structural assessment have included both these ranging technologies. We tested two time-of-flight and two phase-shift scanners for use in the assessment of local topography and the vertical distribution of plant material in a Eucalyptus forest environment in Australia. Plant area index was generally higher when derived using the time-of-flight scanners. The phase-shift scanners also showed less plant area towards the top of the profiles. However, there was a large difference between the plant area derived from the two phase-shift scanner tested, indicating that the ranging technology may not be the primary limitation of such instruments.

1. Introduction

Time-of-flight and phase-shift represent two distinct laser ranging techniques. Time-of-flight ranging employs short pulses of laser energy. Range is determined by measuring the duration between the instrument emitting a pulse and any detected returns. Phase-shift ranging employs constant wave lasers, where the intensity is modulated in a sinusoidal pattern. In this case the phase of any return pulse is used to determine the range of the intercepted object. Both ranging methods have been incorporated into commercial terrestrial laser scanners. In general, time-of-flight scanners have been more commonly employed in vegetation assessment (Cote et al., 2009; Watt and Donoghue, 2005). A number of examples of the application of phase-shift scanners also exist (Antonarakis, 2011; Balduzzi et al., 2011; Park et al., 2010). However limitations introduced by ranging methods used for measuring the vertical distribution of plant area are not well understood.

In this study, four commercially available Terrestrial Laser Scanners (TLS) were evaluated for the derivation of terrain digital elevation models (DEM), the assessment of plant area index (PAI) and vertical profiles of plant area volume density (PAVD). The scanners, listed in Table 1, include two examples of time-of-flight ranging and two that use the phase-shift method.

Table 1: Detail of the laser scanners used in the trial.

Instrument	Faro Focus 3D 120	Leica HDS7000	Leica C10	Riegl VZ1000
Ranging method	Phase	Phase	Time-of-flight	Time-of-flight
Wavelength	905nm	1500nm	532nm	1550nm
Samples/sec	976000	1016000	50000	122000
Beam divergence	0.16 mrad	0.3 mrad	0.1 mrad	0.3 mrad
Nominal max range	120m	187m	300m	1400m

2. Methods

The scanner comparison was undertaken at a research site known as Landers Hut, in the D'Aguilar National Park, Queensland, Australia in November 2011. The site is a Eucalyptus woodland with very little understorey and a ground layer of grass and woody debris. The tree heights were measured within a 50 m by 50 m plot and had a mean height of 13.8 m and mean diameter at breast height (DBH) of 0.18 m. The maximum height within the plot was 37 m and the maximum DBH was 1.05 m.

For each scanner, sampling resolutions were varied to determine if this had an impact on PAI. In addition, range quality, which is associated with measurement integration time, was also varied for the phase-shift scanners. Individual scan details are listed in Table 2. The nominal low resolution equates to the lowest resolution achievable for each instrument and high resolution is the finest angular spacing that could be achieved within 15 minutes of scanning.

Table 2: Angular sampling resolution and phase based scanner range quality setting used in the laser scanner trial.

Scanner	**Nominal Resolution**	**Sampling (mrad)**	**Pulses (x 10^6)**	**Quality**
FARO Focus 3D	High	0.3	178	Low
FARO Focus 3D	Low	3.1	2	High
FARO Focus 3D	Low	3.1	2	Low
Leica HDS 7000	Low	1.2	11	Low
Leica HDS 7000	Low	1.2	11	High
Leica HDS 7000	High	0.3	180	Low
Leica C10	High	0.7	35	NA
Leica C10	Low	1.2	10	NA
Riegl VZ1000	High	0.3	111	NA
Riegl VZ1000	Low	1.2	7	NA

Access to terrain height information is important for TLS data processing as it allows vegetation structure to be analysed in terms of height relative to the ground surface, rather than relative to the origin of the sensor coordinate system. A DEM was generated from each scan using minimum elevation points within a 1m by 1m grid. A progressive morphological filter

(Zhang et al., 2003) was then applied to these points to determine ground returns. Classified ground returns were then used to generate a continuous DEM at a spatial resolution of 1m using the natural neighbour algorithm. Due to the radial sampling of TLS instruments, DEM radius was restricted to 100m from the scan locations.

Each DEM was compared to a DEM generated using Airborne Laser Scanning (ALS) acquired using a Leica ALS50-II sensor with a flying elevation of 1800 m, and recording at a maximum scan angle of 18 degrees off nadir. Pulse repetition frequency of 126.2 kHz produced an average pulse density of 2.8 m^{-2} pulses. The airborne DEM was generated using only returns classified as ground points by the data provider. To be consistent with the TLS data, the airborne DEM was produced at a spatial resolution of 1 m using the natural neighbour algorithm.

Plant density profiles were produced using estimates of canopy gap probability (P_{gap}) in five degree zenith increments between 30 and 70 degrees. Plant density vertical profiles (PAI and PAVD) were calculated from P_{gap} using a linear model described by Jupp et al. (2008). The equation is:

$$- \log(P_{gap}(z,\theta)) \times \cos(\theta) = PAI_h(z) + PAI_v(z) \times 2 \times \tan(\theta) / \pi$$

where z is the height to which PAI is assessed, θ is the zenith angle and PAI_h and PAI_v are the horizontal and vertical projections of the total plant area. Total PAI is then the sum of these horizontal and vertical projections, while the PAVD is the derivative of the PAI profile as follows:

$$PAVD(z) = \delta PAI / \delta z$$

3. Results

The comparison of TLS DEM surfaces to the DEM derived from the airborne lidar data revealed different characteristics between the TLS instruments. While the two Leica and the Riegl scanners showed a gradual positive bias in terrain height relative to the airborne DEM with increasing range from the scanner, the FARO scanner showed significantly lower DEM heights (Figure 1).

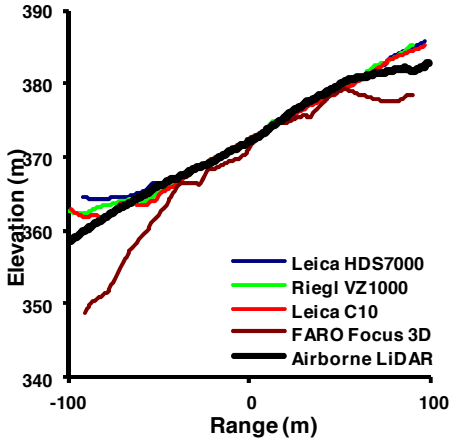


Figure 1: Comparison of Airborne and TLS DEM surfaces along a north-south transect through the Landers Hut site. Airborne Dem is shown in black

Vertical PAVD profiles for the four instruments showed similar characteristics in terms of canopy layering. However, the magnitude of plant density varied considerably, particularly towards the top of the canopy (see Figure 2). The profiles consistently indicated three layers of higher plant area density. While the two Leica and the Riegl scanner showed very stable estimates of PAVD between different scanner settings, the PAVD profiles derived from the FARO data varied considerably depending on the angular sampling resolution and range quality setting used (see Table 2 for details of scan settings used).

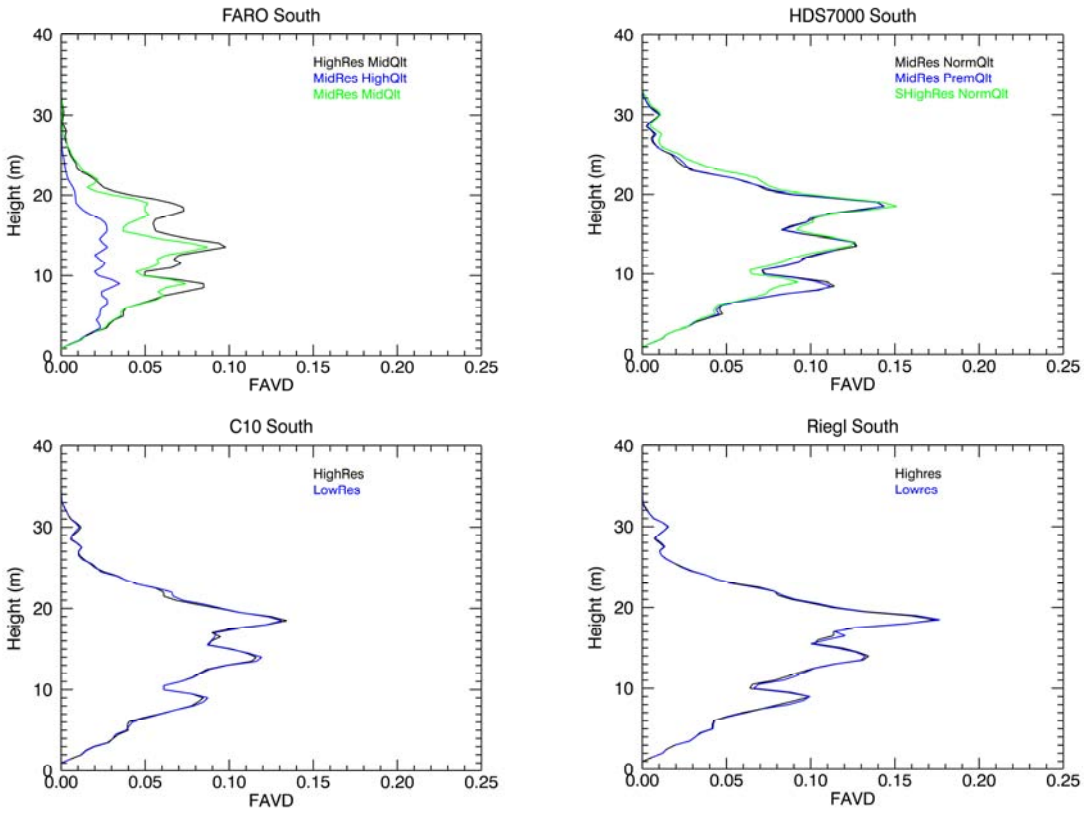


Figure 2: Plant area volume density (PAVD) profiles produced using different setting for each of the four instruments in the south sub-plot. Riegl profiles shown include only first returns in the calculation of P_{gap}. Profiles are not corrected for topography.

The Riegl scanner indicated consistently greater plant density than the other three instruments in the trial. This was particularly true in the dominant overstorey peak at 20m height in the South sub-plot (Figure 2). Although the Leica C10 and Leica HDS 7000 use different ranging technologies, the plant density profiles were remarkably similar, with the HDS indicating slightly higher plant area index. Profiles produced using the FARO Focus 3D showed a high degree of sensitivity to sampling resolution and range measurement integration time.

Total PAI derived from each TLS scan are shown in Table 3 and varied from 0.43 to 2.01. The two Leica instrument were relatively consistent in their estimate of PAI with values varying from 1.70 to 1.82. The Riegl scanner indicated a slightly higher PAI at 1.99 for the high resolution scan and 2.01 for the low resolution scan. The FARO Focus 3D produced the lowest estimates of PAI, varying between 0.43 and 1.15.

Table 3: Plant area index (PAI) derived from each TLS scan recorded at the Landers Hut site.

Scanner	Resolution	Quality	PAI
FARO Focus 3D	High	Mid	1.15
FARO Focus 3D	Mid	High	0.43
FARO Focus 3D	Mid	Mid	0.98
Leica HDS7000	High	Mid	1.82
Leica HDS7000	Mid	High	1.81
Leica HDS7000	Mid	Mid	1.81
Leica C10	High	NA	1.70
Leica C10	Low	NA	1.72
Riegl VZ1000	High	NA	1.99
Riegl VZ1000	Low	NA	2.01

4. Discussion and Conclusions

Our results suggest that inferences about surface topography and vegetation structure derived from terrestrial laser scanners are highly dependent on the scanner employed. Although ranging method (phase-shift verses time-of-flight) may have an impact on data quality, the relatively consistent results achieved using the two Leica instruments suggest this may not be the primary concern.

For the DEM surfaces derived using the TLS instrument in this study, a positive bias relative to the airborne lidar DEM can be explained by the perspective from which a TLS instrument records ranges. Due to the radial sampling pattern there is a gradual increase in occlusion as the ground surface is hidden by near field objects such as trees, woody debris, undergrowth and even the near field terrain its self. The negative bias shown by the FARO instrument is not an artefact of occlusion but is likely due to spurious low points (noise) in the data. This produces depressions in the DEM surface, which are not apparent in the DEM surfaces derived from the other scanners.

Although no validation of the PAVD profiles is provided, the consistency of the PAI produced using the Leica and Riegl scans indicates that the FARO instrument significantly overestimates P_{gap}. The PAVD for the FARO was not only low but was also more biased towards the lower parts of the canopy. Although the nominal range for this instrument is 120 metres, the effective range in a forest environment, where targets are not ideal reflecting surfaces and solar radiant flux is often high, is likely to be much less than that required for characterising the distribution of plant material in the upper parts of a forest canopy.

References

Antonarakis, A.S. (2011). Evaluating forest biometrics obtained from ground lidar in complex riparian forests. Remote Sensing Letters, 2, 61-70

Balduzzi, M.A.F., Van der Zande, D., Stuckens, J., Verstraeten, W.W., & Coppin, P. (2011). The Properties of Terrestrial Laser System Intensity for Measuring Leaf Geometries: A Case Study with Conference Pear Trees (Pyrus Communis). Sensors, 11, 1657-1681

Cote, J.F., Widlowski, J.L., Fournier, R.A., & Verstraete, M.M. (2009). The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. Remote Sensing of Environment, 113, 1067-1081

Jupp, D.L.B., Culvenor, D.S., Lovell, J.L., Newnham, G.J., Strahler, A.H., & Woodcock, C.E. (2008). Estimating forest LAI profiles and structural parameters using a ground-based laser called Echidna. *Tree Physiology*, 29, 171-181

Park, H., Lim, S., Trinder, J., & Turner, R. (2010). 3d Surface Reconstruction of Terrestrial Laser Scanner Data for Forestry. IEEE International Geoscience and Remote Sensing Symposium, 4366-4369

Watt, P.J., & Donoghue, D.N.M. (2005). Measuring forest structure with terrestrial laser scanning. International Journal of Remote Sensing, 26, 1437-1446

Zhang, K., Chen, S., Whitman, D., Shyu, M., Yan, J., & Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne lidar data. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 872 - 882

Forest mapping and monitoring using active 3D remote sensing —Review of a PhD thesis

Mikko Vastaranta[1], Markus Holopainen[1], Juha Hyyppä[2], Mika Karjalainen[2] & Hannu Hyyppä[3,4]
[1]Department of Forest Sciences, University of Helsinki, Finland, first.last@helsinki.fi
[2]Finnish Geodetic Institute, Masala, Finland, first.last@fgi.fi
[3]Aalto University, Research Institute of Measuring and Modelling for the Built Environment, Espoo, Finland, first.last@aalto.fi
[4]Helsinki Metropolia University of Applied Sciences, Finland, first.last@metropolia.fi

Paper Number: SL2012-096

1. Introduction

Forests are mapped and monitored for multiple purposes. In remote sensing (RS), the main aim of mapping and monitoring forests is to produce accurate information for forest managers through the use of efficient methods. In recent years, RS has taken a significant technological leap forward. It has become possible to acquire three-dimensional (3D), spatially accurate information from forest resources using active RS methods. Airborne laser scanning (ALS) has been the most widely used practical application of 3D RS. ALS collects 3D information from forest resources, enabling a highly accurate mapping of tree or stand variables. For example, the estimated root mean square error (RMSE) accuracies for total volume have ranged between 10% and 20% at the stand level in the Nordic countries (Næsset et al. 2004). In addition to ALS, new satellite radars are capable of acquiring spatially accurate 3D information by using radargrammetry or interferometry.

In Finland, the potential applications of RS, such as the utilization of satellite, radar, and aerial images in the spatially accurate mapping of forest variables, has been studied intensively, but the methodologies have not entered general use in practical forest management or pre-harvest planning. The reason is simple: The accuracy in forest variable estimation at the stand level using two-dimensional (2D) RS data has been inadequate. The 3D information obtained from forest resources by active RS has lifted RS-based mapping applications to a new level.

This paper summarizes the main findings of a PhD thesis (Vastaranta 2012) that developed 3D forest mapping and monitoring applications from a Nordic perspective. Mainly original papers are cited, and a broader scope can be found in Vastaranta (2012). The backgrounds of specific mapping and monitoring tasks are presented in chapter 2. Results are reviewed in chapter 3, followed by the conclusion in chapter 4.

2. The backgrounds and main objectives

2.1. Mapping of harvesting sites

Forest mapping and monitoring is carried out to support decision making by the forest owner. In operative forest management planning, input data have been traditionally gathered using stand-wise field inventory (SWFI). In SWFI, wood procurement potential, the amount of round wood removal and forest management proposals are mapped and determined. In addition to stand variables, site types are classified to map forest growth potential, the thinning regime, and

biodiversity. Forest growth and yield are also highly correlated with forest estate value. The wood procurement chain from forest to users starts with knowledge of the stands available for harvesting. The mapping of potential harvesting sites is one of the key decisions for large-scale forest owners (Laamanen and Kangas 2012). The accuracy of the SWFI data has not been adequate for mapping the thinning and final cutting sites. Thus, preharvest measurements have been carried out separately based on existing SWFI causing additional field work.

Currently, the retrieval of stand variables, which is needed in forest management planning, is being replaced by a two-stage procedure using ALS data and field plots, i.e. an area-based approach (ABA, Næsset 2002). Still, forest management proposals are made in the field by foresters. In Vastaranta et al. (2011a) a method to predict stand thinning maturity using ALS data was demonstrated. The method can be used for the mapping of harvesting sites.

2.2 Monitoring of forest canopy changes

In Finland, rather expensive SWFI endeavours have been carried out once every 10 years and updated forest resource information for the intermediate years has been predicted using growth models. Another option is to use continuous updating in forest management, where forest stands are inventoried after each operation and the growth between operations is updated using growth models. However, neither of these methods provides an efficient means to monitor rapid changes in the forests i.e. monitoring applications related to forest growth and the mapping of natural hazards are required at varying scales. Multitemporal, spatially accurate 3D RS data sets are becoming more general, which enables novel monitoring applications. Vastaranta et al. (2012a) present a method for monitoring changes in the forest canopy structure using bitemporal ALS data.

2.3 Efficient forest biomass and volume mapping using ALS and stereo-SAR data

ALS-based inventory methodologies were adopted quickly after the first promising studies (e.g. Nilsson 1996, Næsset 1997a, b, 2002, Hyyppä and Inkinen 1999). Forest inventories using ALS in privately-owned forests were first undertaken in 2010 in Finland, and by the end of 2011, almost 5 million ha had been scanned. Mapping of large forest areas using ALS is usually based on ABA. Mapping could be made more effective by reducing expensive field plots (i.e. reduced amount of field work) or by using inexpensive RS data. Single-tree remote sensing could be used to acquire the modelling data needed in ABA. Vastaranta et al. (2012b) demonstrated a fully RS-based forest mapping method. The method uses ALS data and is capable of producing accurate stand variable estimates even at the sub-compartment level.

Airborne laser scanning is relatively expensive per area unit compared to spaceborne RS data. Thus, other remotely sensed data will still be needed, especially in monitoring applications requiring high temporal resolution. A promising approach to map and monitor forest resources by radar imaging is radargrammetry. Karjalainen et al. (2012) developed a radargrammetry-based method to predict plot-level forest variables as Vastaranta et al. (2012c) compared this approach to state-of-the-art ALS method.

3. Review of the results

3.1 Mapping of harvesting sites

In Vastaranta et al. (2011a), the first tests were conducted to predict the thinning maturity of stands using ABA. The study was carried out in Evo, Southern Finland, and the results were evaluated with 100 test plots located in young and advanced thinning stands. The ground truth

regarding the timing of thinning was determined in the field by an expert. Stands that will reach thinning maturity within the next 10-year period (1) and stands in which thinning should be done immediately (2) were located using logistic regression and k-MSN. Logistic regression models based on ALS point height metrics predicted the thinning maturity with a classification accuracy rate of 79% (1) and 83% (2). The respective percentages were 70% and 86% with *k*-MSN.

Vastaranta et al. (2011a) demonstrated the feasibility of utilizing ALS data for predicting stand-thinning maturity. Subjective expert knowledge was used in the study as a reference and that can be seen as a drawback. On the other hand, it enables one to use it in a wide range of forest conditions and thinning regimes, at least in theory. Närhi et al. (2008) also used ALS features in classifying a stand's precommercial thinning maturity with an overall accuracy of 71.8%. In general, the ALS-based prediction of forest management proposals could provide a practical future means of locating stands with operational needs.

3.2 Monitoring of forest canopy changes

Multitemporal ALS data is of interest in forest monitoring applications. The snow-voluminous winter of 2009-2010 opened up the possibility of studying the use of bi-temporal ALS in snow damage mapping near the Hyytiälä forest research station in Southern Finland. ALS data were acquired in years 2004-2010. The damage was documented in ten permanent Scots pine-dominated plots. Vastaranta et al. (2012a) developed a ΔCHM-method for the detection of snow-damaged crowns. In it, bitemporal ALS CHMs were contrasted and the resulting difference image was analyzed using binary image operations to extract the damaged crowns. Performance was evaluated by errors of omission and commission as well as the error in the estimated damaged crown projection area (DCPA). The method makes use of two threshold parameters, the required height difference (Δh) in the contrasted CHMs and the minimum plausible area of damage (mCC). The best-case performance was evaluated for these parameters and the optimal values were ~1.0 m for Δh and ~5 m^2 for mCC. The plot-level omission error rates were 19-75%, while the commission error rates were 0-21%. The relative estimation accuracy rate of the DCPA was -16.4-5.4%. The strongest predictors of snow damage were stem tapering, relative tree size, and local stand density.

Vastaranta et al. (2012a) had dense, small-footprint ALS data, and the grid size in the CHMs was 0.5 m. Sparser data is likely to be used in practice for detecting corresponding damage. However, it can be assumed that developed method is not oversensitive to the pulse density applied. The average crown size is also an important factor, as it is linked to pulse density. In general, the larger and fewer the crowns are, the less dense the ALS data needs to be.

Vastaranta et al. (2011b) tested the area-based classification of snow damage with multitemporal ALS data. In the study, a forest area was divided into undamaged and damaged grid cells. The predictors used were the ALS point height change metrics, and stepwise logistic regression was used in the classification. The overall classification accuracy for the snow-damaged areas was 78.6% with a Kappa-value of 0.57. Vastaranta et al. (2011b) concluded that area-based estimation is also suitable for snow-induced change detection. Area-based estimation could also detect changes in trees that are not contributing to CHM, which is not possible with methodologies that only use changes in CHMs.

The ΔCHM method is a potential tool for the monitoring of structural canopy changes in the dominant tree layer. Although the method was developed and evaluated in boreal Scots pine-dominated stands, it should be applicable to a wide range of forest conditions with different parameter values. Bitemporal ALS data are not widely available, and the acquisition costs for making a damage inventory would be substantial. Snow damage is a local phenomenon

that is related to topography, while severe storm disturbances occur on a larger scale. Large continuous areas are needed for cost-efficient ALS campaigns, and the methodology proposed here is applicable under such circumstances.

3.3 Efficient forest biomass and volume mapping using ALS and stereo-SAR data

Airborne laser scanning data-based forest mapping applications without any field data

ALS-based forest mapping could be more effective by reducing expensive reference plots (i.e. reduced amount of field work). Vastaranta et al. (2012b) carried out a pilot study by combining laser scanning inventory methods. Individual tree detection (ITD) method from ALS data was used to measure training data for the ABA. In addition to automatic ITD (ITD_{auto}), combination of ITD_{auto} and visual interpretation (ITD_{visual}) was tested. ITD_{visual} had two stages: in the first, ITD_{auto} was carried out, and in the second, the results of the ITD_{auto} were visually corrected by interpreting 3D laser point clouds. ITD_{auto} is usually carried out using only the CHM information, and the understory trees that do not contribute to the CHM are not detected. Visual interpretation is not feasible in mapping but could be utilized when acquiring training data. Vastaranta et al. (2012b) assumed that the human eye can detect understory trees, separate closely growing trees, or drop commission errors easily from the whole point cloud compared with current ITD algorithms.

The RMSE in the imputed VOL was 24.8%, 25.9%, and 27.2% for the ABA trained with field measurements, ITD_{auto}, and ITD_{visual}, respectively. When ITD methods were applied in acquiring training data, the VOL, BA, and Dg were underestimated in the ABA by 2.7-9.2%. Contrary to the assumption, $ABA_{ITDvisual}$ did not provided more accurate results than the $ABA_{ITDauto}$. This phenomenon must relate to the number of nearest neighbours used in the estimation. Absolute accuracy within one field plot is not as crucial when the imputed variable is calculated as a weighted mean over several of the nearest neighbours.

Several other ALS inventory studies have been carried out in the same area. Holopainen et al. (2008) estimated the plot-level VOL with a 27.1% RMSE using 282 field plots for training the *k*-NN method. The pulse density used was 1.8 hits per m^2 and the results were validated using leave-one-out cross-validation. Yu et al. (2010) obtained an RMSE of 56.5% for ITD (without any calibration) and 20.9% for the ABA. Their results were validated with 69 plots, and the pulse density used was 2.6 hits per m^2. In Vastaranta et al. (2012b), a similar level of errors was obtained without any field measurements. However, the pulse density used was much higher (~10 hits/m^2) than those used in the aforementioned studies that favoured ITD. Although the method can be used without any field measurements, it might be practical to use some tree level training data.

The method developed could be applied in areas with sparse road networks or when the costs of fieldwork must be minimized. The method is especially suitable for large-scale biomass or tree volume mapping.

Mapping of forest above-ground biomass and stem volume using TerraSAR-X stereo SAR data

In addition to reducing field data, another option in making RS-based forest mapping more effective is to use inexpensive RS data. Spatially accurate spaceborne SAR data would be suitable for monitoring applications where a high degree of temporal resolution is needed. The results obtained by Perko et al. (2011) showed that X-band stereo SAR satellite data have potential in forest biomass mapping and monitoring even at a substand-level. The use of radargrammetry may also overcome the challenges faced in the mapping of forest variables

using radar-intensity information. TerraSAR-X stereo radargrammetry and sparse nationwide ALS data could be efficient methods for inventorying and monitoring above-ground biomass (AGB) for large forested areas. In radargrammetry, the problem of relating intensity information to forest variables is transformed into a problem of relating the extracted elevation values to forest variables. However, when information about the forest height is obtained, it is a parameter that is highly correlated with forest stem volume and AGB, respectively as in ALS. Karjalainen et al. (2012) developed a radargrammetry-based method to predict plot-level forest variables. 3D points were extracted from stereo SAR images (X-band TerraSAR-X satellite images) to be used as predictors in plot-level forest variable estimation. The extracted point height values appeared to be somewhere between the ground surface and the top of the canopy. The estimation methodologies followed the standard ABA procedures that have been used with ALS data.

The Random Forest method (Breiman 2001) was used in prediction and the relative plot-level RMSEs for AGB and VOL were 29.9%. (41.3 t/ha) and 30.2% (78.1 m^3/ha) when using TerraSAR-X stereo radargrammetry metrics as a predictors. In general, such a high level of prediction accuracy cannot be obtained using spaceborne RS data in the boreal forest zone. For example, Hyyppä et al. (2000) compared SPOT XS, SPOT PAN, Landsat TM, ERS SAR, and JERS SAR data. The relative errors in VOL estimation varied from 45% to 65%. Vastaranta et al. (2012c) compared radargrammetry-based AGB and VOL prediction accuracy to the ALS predictions with the same dataset as in Karjalainen et al. (2012). The ALS prediction accuracies were 21.9% (32.3 t/ha) and 24.8% (64.2 m^3/ha). The ALS imputations were undoubtedly more accurate than the imputations made by using TerraSAR-X stereo radargrammetry metrics. However, the difference between the estimation accuracies of ALS-based and TerraSAR X-based features were smaller than in any previous study in which ALS and different kinds of SAR data have been compared in forest variable prediction (e.g. Hyde et al. 2007, Holopainen et al. 2010). ALS appears to be superior compared to stereo SAR data, mainly due to the much higher point density and lower penetration to the forest canopy. On the other hand, by adding more stereo pairs into the process, the number of 3D points could increase, slightly lowering the relative errors. The future use of spaceborne SAR radargrammetry could be a cost-efficient method for spatially accurate large-area AGB mapping. It should be pointed out that the method requires accurate DEM, which is usually derived using ALS data.

An alternative to radargrammetry when extracting 3D elevation data from radar is interferometry. It has also provided similar level of accuracies in forest variable prediction at the plot level (Solberg et al. 2010a; 2010b).Thus, 3D SAR data appears to be an interesting RS technique for future forest mapping and monitoring. Since SAR satellites enable the mapping of wide areas, they could be used in acquiring detailed forest resource information even at the continental level. The 3D SAR data could also suit forest monitoring well, as the SAR-based point clouds can be adapted to the methods used in operational forest inventories based on ALS data.

4.Conclusions

The possibility of acquiring spatially accurate active 3D RS information instead of 2D has been a major turning point in forest mapping and monitoring. When the aim is to produce as accurate forest resource information as possible for forest managers, this change has opened up totally new possibilities. ALS is an efficient tool for 3D measuring of the forest from above. However, the flying altitudes when acquiring ALS data are relatively low, which makes it expensive per area unit compared to spaceborne RS data. In addition to ALS, other RS data is needed especially if updated information for forest monitoring is required with high temporal resolution. The 3D techniques of interferometry and radargrammetry are appealing spaceborne options in

such cases.

The thesis by Vastaranta (2012) covered a range of applications, and many of the suggested methodologies warrant further studies. Based on the findings in Vastaranta (2012), even fully remotely sensed forest mapping is possible with the same accuracy level as in traditional SWFI. Monitoring of forest biomass changes is one of the near-future applications where active RS data will be required. Data acquisition costs are dropping all the time, and data processing is becoming more automated, enabling its use in larger areas, even in global applications.

Acknowledgements

This study was made possible by financial aid from the Finnish Academy projects 'Improving the Forest Supply Chain by means of Advanced Laser Measurements' (L-impact) and 'Science and Technology Towards Precision Forestry' (PreciseFor).

References

Breiman, L. 2001. Random Forests. *Machine Learning* 45, 5–32.

Holopainen, M., Haapanen, R., Tuominen S. and Viitala, R. 2008. Performance of airborne laser scanning- and aerial photograph-based statistical and textural features in forest variable estimation. Proceedings of SilviLaser 2008: 8th international conference on LiDAR applications in forest assessment and inventory, Heriot-Watt University, Edinburgh, UK, September 17–19, 2008, 105–112.

−, Haapanen, R., Karjalainen, M., Vastaranta, M., Hyyppä, J. Yu, X., Tuominen, S. and Hyyppä, H. 2010d. Comparing accuracy of airborne laser scanning and TerraSAR-X radar images in the estimation of plot-level forest variables. *Remote Sensing*, 2, 432–445.

Hyyppä, J. and Inkinen, M. 1999. Detecting and estimating attributes for single trees using laser scanner. *The Photogrammetric Journal of Finland*, 16, 27–42.

−, Hyyppä, H., Inkinen, M., Engdahl, M., Linko, S. and Zhu, Y-H. 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management*, 128, 109–120.

Hyde, P., Nelson, R., Kimes, D. and Levine, E. 2007. Exploring LiDAR-RaDAR Synergy - Predicting Aboveground Biomass in a Southwestern ponderosa pine forest using LiDAR, SAR, and InSAR. *Remote Sensing of Environment* 106, 1, 28–38.

Karjalainen, M., Kankare, V., Vastaranta, M., Holopainen, M. and Hyyppä, J. 2012. Prediction of plot-level forest variables using TerraSAR-X stereo SAR data. *Remote Sensing of Environment* 117, 338−347. DOI: 10.1016/j.rse.2011.10.008.

Laamanen, R. and Kangas, A. 2012. Large-scale forest owner's information needs in operational planning of timber harvesting – some practical views in Metsähallitus, Finnish state-owned enterprise. *Silva Fennica* 45, 4, 711–727.

Nilsson, M. 1996. Estimation of tree heights and stand volume using airborne lidar system. *Remote Sensing of Environment* 56, 1, 1–7.

Næsset, E., 1997a. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52, 49–56.

−, 1997b. Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, 61, 2, 246–253.

−, 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80, 88–99.

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J. and Maltamo, M. et al., 2004. Laser scanning of forest resources: the Nordic experience, Scandinavian Journal of Forest Research 19, 482–499.

Närhi, M., Maltamo, M., Packalén, P., Peltola, H. and Soimasuo, J. 2008. Kuusen taimikoiden inventointi ja taimikonhoidon kiireellisyyden määrittäminen laserkeilauksen ja metsäsuunnitelmatietojen avulla. *Metsätieteen aikakauskirja,* 1, 5–15.

Perko, R., Raggam, H., Deutscher, J., Gutjahr, K. and Schardt, M. 2011. Forest Assessment Using High Resolution SAR Data in X-Band, *Remote Sensing*, 3, 4, 792–815.

Solberg, S., Astrup, R., Bollandsås, O.M., Næsset, E. and Weydahl, D. J. 2010a. Deriving forest monitoring variables from X-band InSAR SRTM height. *Canadian Journal of Remote Sensing*, 36, 68–79.

−, Astrup, R., Gobakken, T., Næsset, E. and Weydahl, D. J. 2010b. Estimating spruce and pine biomass with interferometric X-band SAR. *Remote Sensing of Environment*, 114, 2353–2360.

Vastaranta, M., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H. and Viitala, R. 2011a. Predicting stand-thinning maturity from airborne laser scanning data. *Scandinavian Journal of Forest Research,* 26, 2, 187−196. DOI: 10.1080/02827581.2010.547870.

−, Korpela, I., Uotila, M., Hovi, A. and Holopainen, M. 2011b. Area-based snow damage classification of forest canopies using bi-temporal lidar data. In Lichti, D. and Habib, A. 2011. LaserScanning 2011 proceedings, pp 5.

−, 2012. Forest mapping and monitoring using active 3D remote sensing. *Dissertationes Forestales* 144. 45p. Available at http:/www.metla.fi/dissertationes/df144.htm.

−, Korpela, I., Uotila, A., Hovi, A. and Holopainen, M. 2012a. Mapping of snow-damaged trees in bi-temporal airborne LiDAR data. *European Journal of Forest Research,* 131 4, 1217−1228. DOI: 10.1007/s10342-011-0593-2.

−, Kankare, V., Holopainen, M., Yu, X., Hyyppä, J. and Hyyppä, H. 2012b. Combination of individual tree detection and area-based approach in imputation of forest variables using airborne laser data. *ISPRS Journal of Photogrammetry and Remote Sensing,* 67, 73−79. DOI: 10.1016/j.isprsjprs.2011.10.006.

−, Holopainen, M., Karjalainen, M., Kankare, V., Hyyppä, J. and Kaasalainen, S. 2012c. TerraSAR-X stereo radargrammetry and airborne scanning LiDAR height metrics in the imputation of forest above-ground biomass and stem volume. Manuscript.

Yu, X., Hyyppä, J., Holopainen, M., Vastaranta, M., 2010. Comparison of Area-Based and Individual Tree-Based Methods for Predicting Plot-Level Forest Attributes. *Remote Sensing, 2*,6, 1481-1495.

Modeling the tree branch structure at very high resolution

Milutin Milenković, Lothar Eysn, Markus Hollaus, Wilfried Karel & Norbert Pfeifer

Vienna University of Technology, Institute of Photogrammetry and Remote Sensing, Gusshausstraße 27-29, 1040 Vienna, Austria; mm@ipf.tuwien.ac.at

Paper Number: SL2012-099

1. Introduction

Knowledge on the structure of vegetation is required in a number of applications. Depending on the scale at which the modeled processes occur or at which the parameters of interest are required, this can either be dominated by larger structures, e.g. a stand or plot, by single tree positions, or at finer details like the individual branching structure. At even finer spatial resolution individual leaves or needles can be described. Modeling as such can occur spatially explicit or as a distribution. In the first case of 3D modeling the available modeling methods use polygons or polyhedral surfaces, freeform curves and surfaces (Farin, 2002), or parametric primitives (Mäntylä, 1998) like, e.g., cylinders or, in the case of planar shapes, ellipses. These geometric elements are assembled in order to represent the geometric structure. Additional information, e.g. optical properties, can be attached to the individual elements. In the case of modeling by distributions, parameters of interest, e.g., the orientation of leaves, are modeled by describing their distribution, e.g. percentiles or average value and variance. Both types of models can – in principle – either be derived from measurements or constructed by assumptions on the structure of the modeled elements.

Radiative transfer modeling using ray tracing allows generating (simulating) synthetic images from a modeled 3D scene by forward modeling (Morsdorf et al, 2012). The scattering behavior within the canopy requires a model of the individual needles or leaves for producing faithful simulation results (e.g. Disney et al., 2010, Disney et al., 2006. Models for extracting the branch structure only from measurements were described, e.g., by Thies et al. (2004) and Bucksch and Lindenbergh (2008). Using only the available measurements of a terrestrial laser scanner (TLS) placed inside the forest and an orchard, respectively, these approaches are not able to model individual needles or leafs. One way to describe for example a forest stand including individual needles or leaves is to use software packages of computer graphics to generate artificially the foliage orientation and variability within the individual tree crowns. For this generation of virtual forest stands several assumptions and generalizations are required. Rutzinger et al. (2010), e.g., extract tree location, height, stem diameter and crown diameter from mobile laser scanning data and "grow" tree models (Weber and Penn, 1995) reaching the measured parameters. Côté et al. (2009) used for the reconstruction of 3D tree architecture from terrestrial LiDAR scans a generic conifer shoot model based on the model description of Smolander and Stenberg (2003). However, there is still a lack of detailed information concerning actual shoot structure as stated in Côté et al. (2009).

To overcome the limitation of actual shoot structure information, we present an approach for reconstructing exact geometric information on individual twigs, needles, and leaves from measurements. This reconstruction can be performed with sub-millimeter accuracy using a measuring arm in an indoor setting. At the current stage, the approach requires, next to the measurement itself, manual interaction. Individual needles and leaves as well as twig fragments are modeled in terms of their position, length and direction. Basics of the applied 3D modeling approaches were developed within the NEWFOR (Hollaus, 2012) project for a coarser scale (i.e. tree structure) and are adapted and improved for branch modeling within the 3D-VegetationLab project.

In Sec. 2 the measurement device and the sample branches, one coniferous and one deciduous,

are described. In Sec. 3 the methods are explained and the results are discussed in Sec. 4.

2. Study objects and data

2.1 Study objects

One coniferous and one deciduous branch are selected as objects for the modeling. Both branches belong to the end part of the corresponding tree limbs, and carry a typical structure of their species, which makes them – in principle – also suitable for cloning all over the tree. The objects also have similar structures in the sense that both branches are composed of one "primary" and several "secondary" branches.
The coniferous branch (Fig. 1) is taken from a fir tree (*Abies alba*) and is approximately 80 cm long and completely covered with needles. The "secondary" branches spread about 40 cm sideways from the "primary" one, mainly in one plane. They contain many shoots that are attached to them. While not of importance for the reconstruction method, the oldest part of the branch was estimated to an age of 5 years. The shoot location within a secondary branch was used to determine the age of the shoot, with the end shoots considered as first year shoots, whereas others decrease in the age while moving from the branch end towards the branch joint.
The deciduous branch (Fig. 2) comes from a European beech tree (*Fagus sylvatica*). Compering to the coniferous branch, it is slightly bigger in size (c.a. 1 m in length), but has a similar planar spreading of the secondary branches. Twigs with attached leaves are sparsely distributed over the secondary branches making the object structure more "open" and less complex. Scanning of the branch indoors required a fast transportation of the cut branch to the measurement device because of the fast drying out process.

2.2 Measuring device

A measuring arm (METRIS MCA, 3600 M7) is used to acquire exact geometric information. The instrument used has a triangulation laser scanner mounted on its head, and therefore allows contactless object scanning. A laser stripe is constantly emitted and intersects with an object's surface if pointed at it. This illuminated cross section is mapped by a camera and furthermore 3D coordinates of the illuminated area are derived. If the laser stripe is moved over the object, a so called "scanning strip" is recorded. Larger areas are covered by scanning several strips. The METRIS measuring arm provides automatically oriented scans of the object. The mounted scanner provides a point cloud that has different along- and across-strip resolution. The across-strip resolution depends upon both the laser strip width and the resolution of the camera, and goes up to 0.05 mm. The along-strip resolution, on the other hand, is usually about 0.5 mm, which is mainly driven by the moving speed of the arm's scanning head. A slow movement of the head can significantly increase this value and provide a large number of acquired points. Due to its construction, the instrument has a limited operational range that allows scanning of objects up to 1.5 m in size. Besides, there is an operational limit which dictates the size of the acquired scan, i.e. the number of points that can be processed by the system.

2.3 Acquisition method

The arm's resolution and operational range, on the one hand, and the objects' structure and size, on the other hand, appear as the most prominent factors during the scanning. Those specific characteristics of the instrument and objects are the reason for introducing separate acquisition strategies for coniferous and deciduous branches.

2.3.1 Coniferous branch

The complex structure of the conifer branch and its large extend (compering to the operational range) affect both object stability and accessibility during the scanning procedure. Therefore, a two-stepped strategy is introduced to acquire a representative point cloud of the model. First, the overall structure is in the focus, then, each "secondary" branch is measured. In the first step, the main branch and large secondary branches are scanned using a number of scan strips and

ensuring overlaps between the scans. Then, the secondary branches are separated physically (i.e. cut) from the object and measured separately. Therefore, co-registration of all the scanned parts is required in the second step (Fig. 1).
The co-registration is done for each secondary branch independently, transforming its coordinate systems to the main branch's coordinate system. This procedure is done using the spherical heads of pins as targets that are present in both scans. Tree pins per each secondary branch are introduced in zones close to branch joints i.e. in the overlapping scan areas (Fig. 1). Based on the points that represent those targets, different spheres are fitted and then their centers are used for the co-registration. This acquisition approach also allows to model each secondary branch independently and latter to build the final model by merging all the models into one structure.

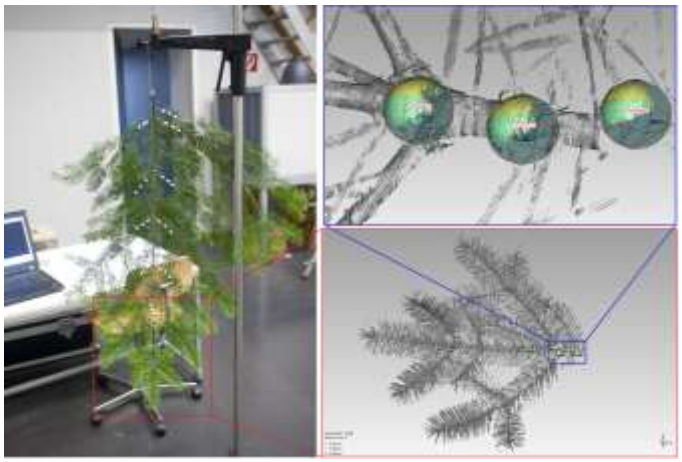


Figure 1: Left: conifer branch prepared for the measurement; Bottom right: detail showing the point cloud of a scanned ("secondary") part; Top right: detail showing the spheres for co-registration.

The point cloud of the coniferous branch (Fig. 2, right) has over 40 million points. The branches are covered from both sides. As visible in Fig. 1 (bottom right), the first and second year shoot at the end of the tree are not perfectly symmetrical. This holds for the directions and lengths as well as for the branching structure.

2.3.2 Deciduous branch

The deciduous branch is scanned in one step because its sparse structure allows an easy access to the entire object. However, the area of the leaves and the high resolution of the scanner produce a large amount of data that exceeds the operational limit of the arm. Therefore, several smaller scans (i.e. files) are collected to cover the whole object. Since the object and scanner coordinate system where fixed during the scanning, there was no need for the final co-registration of these scans. The final point cloud of the deciduous branch is shown in Fig. 2 (left). It contains almost 50 Million points. The scans were only acquired from the leaf top and bottom side.

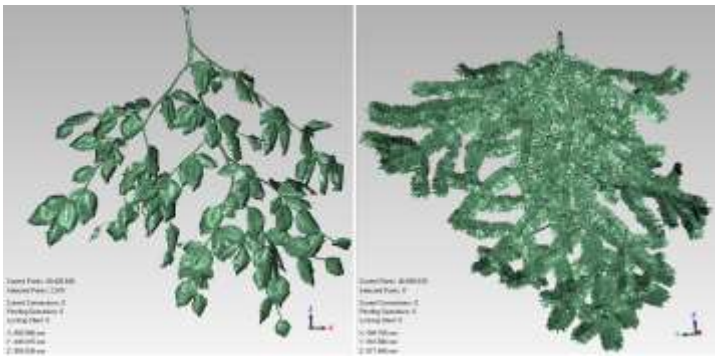


Figure 2: Left: Point cloud of the deciduous branch; Right: Point cloud of the coniferous branch

3. Methods

Needles and leaves are the elementary unit in modeling the coniferous and deciduous branch, respectively. Since it is not possible to represent them in a geometrically similar way, two separate methods are proposed to model those objects. Both methods start with a common task where the corresponding branch structures are modeled, and then, during the modeling of their elementary unit, they split into two independent procedures. Following this logic, the Sec. 3.1 will first discuses the extraction of the branch skeleton, and in Sec. 3.2 and Sec. 3.3 the proposed modeling methods are described.

3.1 Modeling branch structure

The branch structures of the two objects are very similar, i.e. both have branch joints and cylindrical wooden parts that connect them. Thus, topologically correct 3D polylines are used as a model for both the coniferous and deciduous branch structure. Those polylines are digitized (manually) from the corresponding point clouds.

The open structure of the deciduous branch allows a fast digitization without interruptions of the entire object structure in one coordinate system, i.e. the object coordinate system (OCS). This included digitization of the leaf stalks.

The digitization of the coniferous branch structure is not so straightforward. The presence of needles disturbs the digitization process making the selection of stem points much more time consuming compering to the deciduous branch. Therefore, an automatic procedure for point reduction is implemented before the digitization. The point cloud reduction is based on the voxel model derived from the scanned part and an estimation of a local voxel density. A property of the original point density is, that it is strongly influenced by the instrument's scanning pattern.

Voxelizing the data with an appropriate voxel size, the native structure of the object is emphasized rather than the scanning properties of the instrument. A voxel is considered as foreground (i.e. filled), if it contains at least one point from the original point cloud. The voxel density is the percentage of filled voxels in the 33-neighborhood of the center voxel (sphere of radius 2 in voxel space). An empirically driven thresholding on the local voxel density is done to reduce the starting point cloud. The used voxel size is 1 mm, which is close to the width of an average needle of the conifer branch. After the reduction procedure most of the needle points are removed from the starting data set, and the remaining points (mainly points from the stem) are used for digitizing the skeleton of the scanned part (Fig. 3). Since the co-registration of all the scanned parts is already done, the resulting 3D polyline structure is in a coordinate system of the conifer branch, i.e. in the OCS.

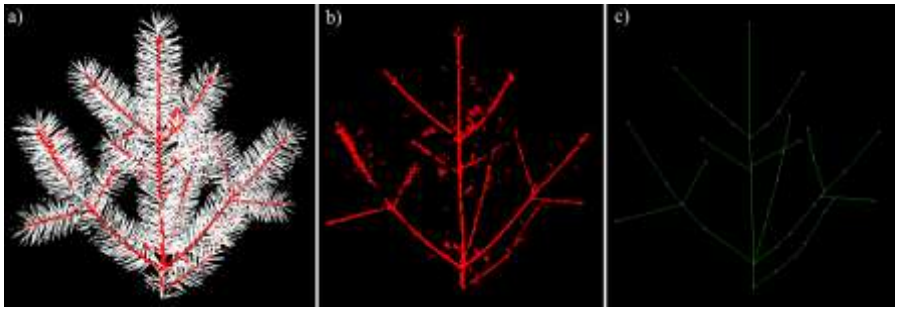


Figure 3: a) The classified original point cloud, based the thresholding the local voxel density; b) the reduced point cloud; c) the skeleton of the scanned part (right)

3.2 Modeling the deciduous branch

For modeling the deciduous branch it is assumed, that the wooden structure is available as 3D polyline. Thus, only the leaves need to be modeled. For each leaf, a local point cloud is extracted. This is based on a bounding box which is extracted per leaf. Based on those local

point clouds, for each leaf an ellipse is estimated and used as a model for that leaf. Before estimating the ellipse, the regression plane is fitted trough the leaf point cloud (Fig. 4, left). This provides the orientation of the leaf plane.
Next, the leaf point cloud is orthogonally projected to the regression plane and boundary leaf points are extracted from this data set. The ellipse parameters (center point, major and minor radius, and axis orientation) are estimated by fitting the ellipse trough the leaf boundary points (see Figure 4, right). These parameters are transformed into a description suitable for including it in the model coordinate system (MCS): the normal vector of the regression plane, the ellipse center point, the end point of the longer radius of the ellipse, and the minor radius.
To establish a correct topology in the final model, the estimated ellipse is translated to the closest end point of the branch skeleton. In this way the native orientation of the leaf (the regression plane) is preserved. The translation parameters are resolved from the equality condition of the end point of the longer radius of the ellipse (tip of the leaf) in MCS and the closest end point of the branch skeleton in OCS.

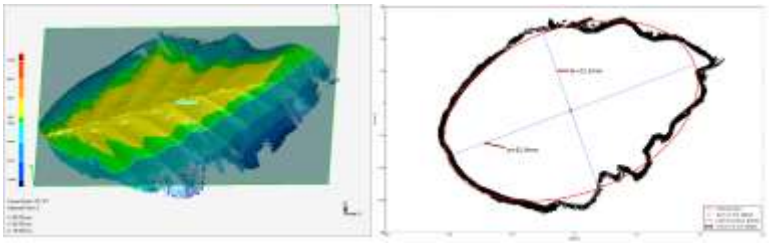


Figure 4: Left: The leaf point cloud with the regression plane; Right: the fitted ellipse displayed over the leaf boundary points

3.3 Modeling the coniferous branch

To handle the large data set, the modeling of all the scanned parts (the primary and the secondary branches) is performed independently and then, in the subsequent step, their models are combined into the final branch model. In the following a method for modeling of a shoot is presented. This method can either be applied to the entire branch, or the shoot can be cloned to populate the entire skeleton with the needles of the shoot model.
Using cloning is therefore a two step approach: (1) creating the shoot model, and (2) cloning the shoot model over the branch structure. A detailed workflow that includes also the skeleton modeling and all intermediate layers is illustrated in Figure 5. Since the proposed strategy allows an independent modeling of the scanned parts, only one is selected here to demonstrate the potential of the proposed method. The end part of the branch (see Figure 1) is taken because it has the most complex structure among other parts. Using a shoot model multiple times, enhances the efficiency of the modeling procedure. However, a typical data set needs to be selected when creating only one shoot model. This 3D model is assumed to be representative for all the shoots within the branch and used later, in the cloning step.

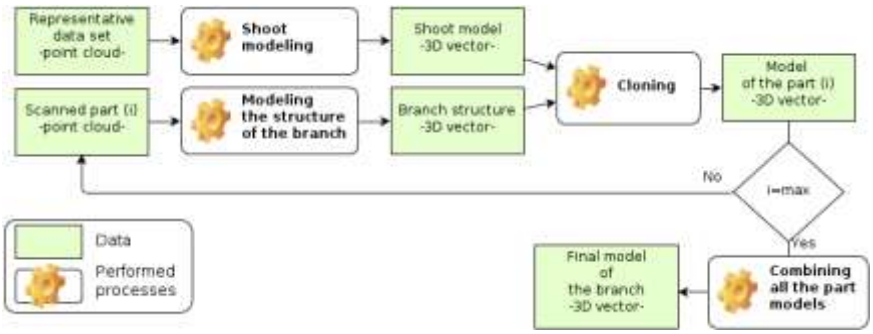


Figure 5: The workflow for modeling the branch

3.3.1 Modeling the shoot

A point cloud of a first year shoot (Fig. 6a) is selected as a representative data set for the shoot modeling by manually digitizing. Each needle is represented by a 3D line, with a starting and ending point, while the shoot's stem is represented by 3D polyline. Since the starting points of the needle lines are also the nodes of the stem poly line, the model is also topologically correct (Fig. 6b). The local coordinate system of the shoot model - the model coordinate system (MCS) - is defined with two reference points and one plane. The starting and ending point of the wooden part are selected as the reference points, whereas the regression plane fitted trough the needle end points (Fig. 6c) is taken as the reference plane.

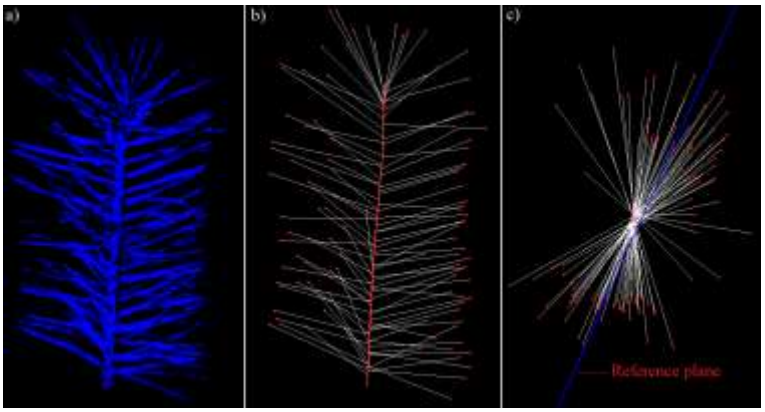


Figure 6: a) Point cloud of the representative shoot; b) the shoot model; c) the shoot model from the top view with the reference plane.

3.3.2 Cloning the shoot model

To build the conifer branch model by cloning, the shoot model is cloned over the branch skeleton. For each segment of the skeleton's polylines an appropriate part of the shoot model is firstly extracted and then cloned on this location. For example, segments that represent an end part of the skeleton are cloned with the whole shoot model, whereas middle segments are cloned only with a lower part of the shoot model. In case where the length of the skeleton segment exceeds the length of the shoot model, the lover part of the shoot model is used to bridge the remaining part of the segment.
The cloning procedure is performed by transforming the shoot model from MCS to OCS. Since the MCS is defined by the two points and one plane, the starting and ending point of the skeleton segment are used as the reference points, while a local plane is taken to resolve the orientation of the shoot model. The local plane is estimated for each skeleton segment independently, based on the local needle points filtered out during the point reduction procedure (see Sec. 3.1).

4. Results and discussion

In this Section the results of the cloning process are presented and discussed. The results are based on the shoot model shown in Fig. 6b, and the branch skeleton shown in Fig. 3c. The application of the above described modeling and cloning methods lead to the branch model shown in Fig. 7. The point cloud and the final branch model show a perfect correspondence for the location of the manually modeled shoot model (marked red in Fig. 7), exhibiting the same (ir)regularity of the needles especially at the end and with respect to their orientation out of the plane. The lower part of the shoot model shows a homogeneous density of needles. Therefore, this 1st year shoot could be cloned all over the manually extracted branch skeleton. At joints, where multiple sub branches connect, the 3D model, in some cases, exhibits more needles than the point cloud data suggests (Fig. 7).

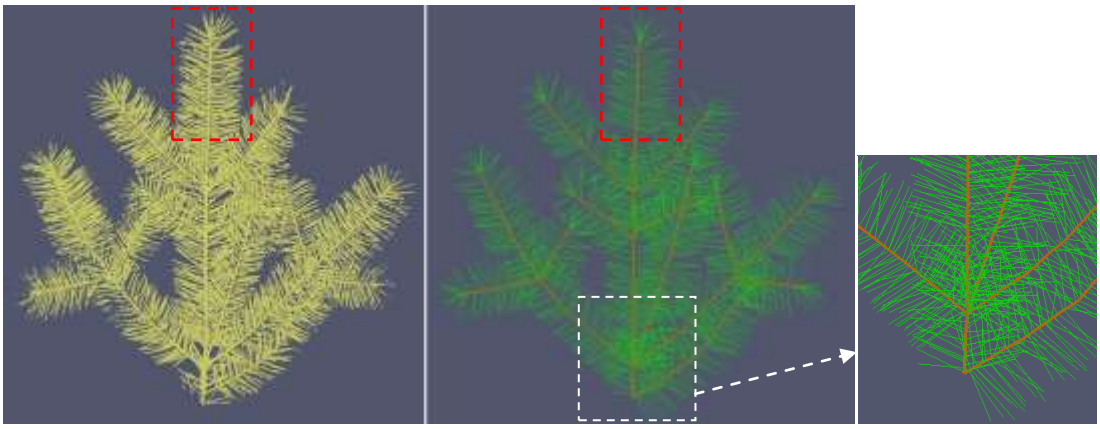


Figure 7: Point cloud of a coniferous twig versus it's extracted 3D model. The red boxes show the shoot which was manually modeled and used for cloning; Left: Pointcloud; Middle: 3D model with classification into needles and branch structure; Right: Erroneous clumping of needles at joints

Considering the entire method of modeling, a number of steps include simplifications and cause differences between the geometric model and the acquired data. However, the entire method is data driven and allows modeling of the branch structure for each individual needle and leaf as well as the branches in between. A slight curvature of the needles was visually found in the data set but was neglected during modeling. The chosen model for a needle is the straight connection between the start and the end point of the needle. This rather simple model could be enhanced by using the average thickness of the needle for fitting an ellipsoid of revolution, which is favorable if the models are used i.e. for ray tracing analyzes where volumetric objects are preferred. However, also the orientation of the branch is available, which allows even more complex needle models to be applied, too.

Because the branch skeleton and the needle axis are digitized manually based on the point cloud, the location of the digitized elements is on the outer surface of the corresponding objects. If necessary, this error could be simply corrected by estimating or measuring the diameter of the corresponding object and translating the axis by half of the diameter. While a fitting approach would estimate the axis directly, the irregular point distribution, especially, in the case of a coniferous branch, poses an additional challenge to obtain a better geometric position in this way.

Concerning cloning, no 2^{nd} year, 3^{rd} year or even older shoots were used for this model. In general, the 1^{st} year shoots have more needles than the older shoots and therefore an overestimation of the number of needles in the current cloned model is assumed. 2^{nd} or 3^{rd} year shoots could be simulated by reducing the number of needles before inserting the clone into the branch structure. Further investigations in this direction need to be performed.

Concerning the leaf model, the approximation of the leave as an ellipse applies well to the beech leaves used in this study. However, the main consideration is a model with a small number of parameters, e.g. 5 in the case of an ellipse or 6 for a triangle. Fitting elements like an ellipse overcomes problems with gaps or inhomogeneous point densities in the acquired datasets. Alternative approaches include using the alpha-shape for the outline detection and the Douglas-Peucker algorithm for simplification in order to get a polygonal outline with a small number of points. Generally, the leaves are assumed to be flat which does not account for seasonality, although leave curvature is visible in Fig. 4 (left). Again, ray tracing advocated for simple models.

5. Conclusion

Methods to model individual twigs, needles, and leaves, based on point clouds acquired with a measuring arm, at a high level of detail are presented. The model for coniferous branches

consists of 3D polylines with an additional classification, whether a line element represents a needle or belongs to the branch. Each line element can be transformed to a volumetric description by using their respective thickness value. The chosen model for representing a single leaf is an ellipse which acts as a good approximation for the complex leave structure. This study showed that the presented approach leads to promising results, which represent the real 3D branch structure based on measurements of the object. These detailed coniferous and deciduous 3D branch models can be used to build up larger twigs and branches by using cloning approaches. In this way virtual tree models can be enhanced or created that are used i.e. as input for radiative transfer and process modeling. While a faithful reconstruction is possible with the presented approach, measures of quality need to be derived before these results can be used in further applications. Likewise, for larger scenes to be modeled, the degree of automation needs to be increased.

Acknowledgements

This study is partly financed by the ESA funded 3D-VegetationLab project and the project NEWFOR financed by the European Territorial Cooperation “Alpine Space”.

References

Bucksch, A., Lindenbergh, R. 2008. Campino – a skeletonization method for point cloud processing. ISPRS Journal of Photogrammetry and Remote Sensing 63(1), 115-127.

Côté, J.-F., Widlowski, J.-L., Fournier, R. A., and Verstraete, M. M. 2009. The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. *Remote Sensing of Environment*, 113(5), 1067-1081. Elsevier Inc. doi:10.1016/j.rse.2009.01.017

Disney, M., Lewis, P., and Saich, P. 2006. 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains. Remote Sensing of Environment 100(1).

Disney, M., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S., Pfeifer, M. 2010. Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. Remote Sensing of Environment 114 (7), 1546-1560.

Farin, G. 2002. Curves and surfaces in CAGD. 5th ed. Academic Press.

Hollaus, M., Mücke, W., Eysn, L. 2012. Forest Structure And Stem Volume Assessment Based On Airborne Laser Scanning. AMBIÊNCIA (ISSN 2175-9405), accepted.

Mäntlyä, M. 1998. An introduction to solid modeling. Computer Science Press.

Morsdorf F. and 9 co-authors 2012. 3D-Vegetation Laboratory: A scientific support tool for accuracy assessment and prototyping of EO data and products. Proceedings of ForestSat 2012.

Rutzinger, M., Pratihast, A.K., Oude Elberink, S., Vosselman, G., 2010. Detection and modelling of 3d trees from mobile laser scanning data. Int. Archives of Photogrammetry and Remote Sensing, Vol. XXXVIII/5, Newcastle upon Tyne. 520-525.

Smolander, S., and Stenberg, P. 2003. A method to account for shoot scale clumping in coniferous canopy reflectance models. *Remote Sensing of Environment*, 88, 363-373. doi:10.1016/j.rse.2003.06.003

Thies, M., Pfeifer, N., Winterhalder, D., and Gorte B.G.H. 2004. Three-dimensional reconstruction of stems for assessment of taper, sweep and lean based on laser scanning of standing trees. Sca ndinavian Journal of Forest Research 19 (6), 571-581.

Weber, J, and Penn, J. 1995. Creation and rendering of realistic trees. ACM International Conference on Computer Graphics and Interactive Techniques, New York, United States, pp. 119-128.

Mapping and monitoring urban trees using airborne scanning LiDAR and tree register data

Topi Tanhuanpää[1], Mikko Vastaranta[1], Markus Holopainen[1], Ville Kankare[1], Juha Raisio[2], Petteri Alho[3], Juha Hyyppä[4] & Hannu Hyyppä[5]

[1]University of Helsinki, Department of Forest Resource Management, Latokartanonkaari 7, 00014 Helsinki, Finland, (topi.tanhuanpaa, mikko.vastaranta, markus.holopainen)@helsinki.fi)
[2] City of Helsinki, Public Works Department, Street and Park Division, Kasarmikatu 21, P.O. Box 1515, 00099 Helsingin kaupunki,Finland (juha.raisio@hel.fi)
[3] Department of Geography and Geology, University of Turku, FI-20014 Turku, Finland (mipeal@utu.fi)
[4] Finnish Geodetic Institute, Geodeetinrinne 2, PL 15, 02431 Masala, Finland. (juha.hyyppa)@fgi.fi)
[5] Aalto University, Research Institute of Modelling and Measuring for the Built Environment, Finland (hannu.hyyppa@aalto.fi)

Paper Number: ### SL2012-103

1. Introduction

Urban trees and forests have many functions that differ from trees in commercially managed forests (Bolund and Hunhammar 1999). Benefits such as noise reduction and recreation affect the life of all the people living near urban forests or trees. This creates both commercial and non-commercial value for urban trees. During recent years carbon sequestration and biomass of urban trees have also become more interesting. Although urban trees are rarely a source of financial income, planting and especially nursing of the planted trees are rather costly tasks. Planting and nursing actions can be arranged more efficiently if data in tree registers is up-to-date. Among many other cities, the city of Helsinki maintains a large register of its roadside and park trees. The tree register contains information about the locations and characteristics of individual trees. The register is updated approximately once in every ten years, which is too infrequent for many planning purposes. In a dynamic urban environment, out-dated information hinders urban planning. Especially roadside trees are more likely to encounter radical changes in their surroundings than trees in park areas or forests. The quality of the register varies in both tree location and characteristics. Variation of quality in tree location data arises from different localizing methods, whereas the quality of other tree characteristics depends mainly on the age of the information. Trees have been located either visually – using maps and aerial photographs – or using GPS measurements. All tree characteristics have been measured manually in the field. Now the city of Helsinki is seeking cost efficient methods for urban tree mapping and maintaining accurate tree registers.

Remote sensing (RS) provides efficient tools for urban tree mapping and monitoring, especially airborne scanning LiDAR (Light Detection and Ranging). With LiDAR, it is possible to collect dense 3D point clouds that describe the structure of the mapped object, whether it is terrain, building or tree. Up-to-date high density (>10 XYZ-points/m^2) airborne scanning LiDAR data is widely available as it is used for multiple planning purposes in urban engineering. Hence urban environment enables the use of datasets that would normally be too costly, at least for non-commercially managed trees. The use of LiDAR-based forest mapping has been widely studied (e.g. Næsset et al 2004, Hyyppä et al. 2008, McRoberts et al. 2010). There are two separate methodologies in utilizing LiDAR data in order to predict forest and tree characteristics: Area-

based Approach (ABA, Næsset 2002) that is based on statistical dependence between field plots and LiDAR metrics (i.e. features calculated from 3D points) and Individual Tree Detection (ITD, Hyyppä and Inkinen 1999) that recognizes individual trees from the more or less processed LiDAR data. In ITD, characteristics of individual trees can also be predicted statistically from 3D metrics. The precise nature of data needed in urban planning as well as mapping of trees and tree characteristics suggests that inventory should be done at tree level, i.e. ITD is likely to be well-suited for mapping applications in urban areas. The method defines the location and characteristics of an individual tree such as height, diameter and volume (Hyyppä and Inkinen 1999, Holmgren and Persson 2004, Popescu et. al. 2004, Koch et. al. 2006, Falkowski et. al 2008, Maltamo et. al. 2009). When compared to purely manual field measurements, LiDAR-based forest mapping applications have proven to be cost-efficient and reliable in commercially managed forests (Holopainen and Talvitie 2006, Uuttera et. al. 2006). The main advantage of LiDAR methods is decreased requirement for costly field measurements. LiDAR-based tree mapping methods have also been implemented in urban environment although the amount of publications is limited. The focus of urban studies has usually been set to either automatic segmentation of individual trees (Palenichka and Zaremba 2007) or estimation of tree characteristics (Shrestha and Wynne 2012). Omasa et al. (2008) produced detailed 3D-models of park trees by combining airborne and terrestrial LiDAR-data. There are also a few companies that offer LiDAR-based mapping of urban trees as a commercial product, but accuracies and details of the methods used are not public.

In this study, the use of high density airborne scanning LiDAR data was studied in mapping and monitoring of the key attributes of urban trees. Our aim was to create an effective method for updating an existing urban tree register. This included mapping of trees using ITD, modelling their characteristics and combining updated data with the existing tree register.

2. Materials and methods

2.1 Study area and existing tree register

The study area consists of roadside areas in central and western part of Helsinki, including the city centre. In total, there were approximately 10 000 roadside trees in the study area, and they represented over 60 different tree species. However, vast majority (75 %) consisted of three deciduous tree species: lime (Tilia sp.) (55 %), maple (Acer sp.) (11 %) and birch (Betula sp.) (9 %).

The City of Helsinki supplied an excerpt from the city's tree register that contained all roadside trees in the study area. The most essential tree characteristics were the location and species of each tree. The age of the tree data in the register varied from 1 to 12 years. However, majority of the register trees had last been updated in 1999, which leads to data that is 12 years old. The age of the data meant that all trees had grown to some extent and some of the trees had been removed or replaced. There were also defects in the register, such as significant location errors, mainly caused by mistakes in the maintenance. The tree register has been both created and updated manually, which is why we call it the *manual tree register*. A mixture of methods used in creating and updating the register has led to errors of varying degree, especially in the accuracy of the location of the trees.

2.2 Field measurements

The field data was measured in autumn 2011. The aim of the measurements was to collect reference data to update an existing tree register, and the following procedure was carried out. Sample trees were chosen in such a manner that the diameter and species distributions of the measured trees resembled the whole of the roadside tree register. Many of the 60 species groups consisted of only a small number of trees. Therefore, the smallest groups were combined with

other groups to form larger strata. In total, 14 strata were used, and the field-measured trees represented 22 different tree species altogether.

Sample trees were measured streetwise to enhance the field measurement efficiency: areas consisting of desired tree species and diameter classes were sought for. Once found, the whole street section was measured. Diameter at breast height (DBH) was collected from 1239 trees. Tree height was measured from 576 sample trees. Measured area covered roughly 9 km of road sections.

2.3 Remote sensing data

Remote sensing (RS) data consisted of high density airborne LiDAR data and high-resolution aerial photographs. The study area was scanned at spring time in 2009 and 2010. Scanning data sets were so-called "leaf-off" data which is desirable in urban planning purposes. Leaf-on data is more common in forest mapping applications. The density of returned LiDAR pulses was about 20 points per m^2. In addition to the LiDAR data, same-date high-resolution aerial photographs were used in the study as visual reference. The resolution of the images was approximately 5 cm.

2.4 Urban tree mapping using individual tree detection

Urban trees were mapped using ITD procedures, the accuracies of which are well known in forest environment (Vastaranta et al. 2010, Vauhkonen et al. 2011, Kaartinen et al. 2012). Readers are advised to see original references for more detailed description of ITD (Hyyppä and Inkinen 1999, Popescu et al. 2004, Holmgren and Persson 2004, Yu et al. 2011), as it is only briefly described here. The used tree delineation method was based on the creation of normalized surface model (nDSM, or in the case of a forest, canopy height model or CHM), and individual trees were then delineated using watershed segmentation from smoothened nDSM (e.g. Hyyppä and Inkinen 1999, Holmgren and Persson 2004). The urban environment set special challenges for the tree extraction process. Although roadside trees are relatively sparsely located, there is a lot of urban infrastructure that complicates the tree extraction process: lamp posts and power lines, for example, reflect the laser pulses and affect the DSM/CHM. In addition to infrastructure, high vehicles may also affect the segmentation process. Segments that represented trees had to be separated from the irrelevant segments (e.g. lamp posts, etc.). After tree delineation, LiDAR point clouds were extracted for each segment (i.e. tree candidate). Every segment was considered to contain only one tree. In total, 23 LiDAR metrics were calculated. The metrics included only the most common LiDAR metrics, such as height percentiles and propositions of canopy hits at various heights. For more detailed description of the used features, see e.g. Yu et al. 2011. Only "first" and "only" echoes were used in these calculations.

2.5 Linking ITD tree candidates with existing tree register

After extraction, the segments were joined to corresponding trees in the tree register. At this point, the locations of the trees in the original tree register were utilized to join the tree segments and the register trees. When joining the tree candidates and the location points of register trees, the segments were divided in three groups:

1. Segments with only one register tree location point in them
2. Segments with multiple register tree location points
3. Segments with no register tree location points

Segments in the first group were accepted as they were and the two latter two groups were further examined. The first tree group took in about 90 % of the trees in the tree register. These segments were combined to register trees based on XY-location.

The second group consisted of segments that included multiple register trees. This kind of segmentation occurred if the register trees had overlapping crown area or the positions of the trees were incorrect. It is also possible that the segments in the second group contained so-called "ghost trees" that no longer existed. Such trees had died or had been cut down at some point, but their entries had not been erased from the register. This group was the most laborious to work with, because all the segments had to be manually checked for “ghosts”. This inspection was carried out with visual interpretation of high resolution aerial photographs and LiDAR point clouds. The segments in the second group that did actually contain multiple trees were ignored and new segments were created manually for each register tree.

The segments in the third group were tested for nearby trees: there were erroneous tree locations up to several meters in the original tree register. Therefore, the segments in the third group were joined to the nearest register tree. However, this could be done only when the first two groups had been completed.

2.6 Prediction of tree characteristics

The field-observed tree variables, i.e. height and DBH, were used as Y-variables and the LiDAR-derived metrics as X-variables. Prediction was carried out using the Random Forests (RF) method which can be considered as one of the NN (Nearest Neighbour) techniques. The main idea in NN imputations is that two records which share similar X-values should also have similar Y-values. Then the missing Y-variables for target records can be imputed using some similarity metrics and the reference records with both X- and Y-variables. In RF, two records are considered similar if they tend to end up in the same terminal nodes in a collection of classification and regression trees (Breiman 2001). The distance measure in RF is one minus the proportion of trees where a target observation is in the same terminal node as a reference observation (Crookston & Finley 2007).

In the RF method, 500 trees per imputed variable were fitted in each RF run and square root of the number of predictor variables was picked randomly at the nodes of a regression tree. In addition to the number of fitted trees, the number of neighbours (k) that is used in predicting the variables is an essential parameter affecting the results. In this study, k-values from 1 to 15 were tested. 100 RF runs were executed for each k and root-mean-square deviation (RMSD) was determined for each tree. When k-value was set to 7, the method led to the smallest RMSD. However the difference in RMSD between k-values of 5, 6 and 7 was so small that value 5 was chosen as the final number of neighbours. The decision of using smaller k-value although its RMSD was slightly higher was made because bias increases as the number of neighbors increases. Selecting the value for k is always a compromise: low value of k increases the random error of the estimates, while high value of k results in averaged estimates and reduces the variation available in the original dataset. The R yaImpute library (Crookston & Finley, 2007) was applied in the imputations.

3. Results

Tree variables were estimated for 1239 field trees. The estimated variables were DBH, tree height and location of the tree (X- and Y-coordinates). The error estimates of height and DBH estimations are presented in table 1.

Table 1: RMSE of tree height and DBH by measurement unit

Measurement unit	$RMSE_{abs}$		$RMSE_{rel}$		Relative size of the measurement unit (%)
	height (dm)	DBH (cm)	height (%)	DBH (%)	
Lime	10,4	4,3	10,1	16,0	47
Rowan	10,7	3,8	16,0	21,4	16
Elm	5,9	4,8	5,9	19,8	7
Maple	17,0	6,1	14,0	25,9	6
Birch	6,1	5,1	4,3	24,6	5
Oak	3,5	2,4	4,8	15,2	3
Apple	8,9	6,5	11,5	31,9	3
Pillar trees	7,1	3,6	6,1	19,1	3
Hawthorn	5,6	1,8	8,8	12,7	3
Alder	12,8	4,9	7,5	15,0	2
Willow	15,6	18,8	10,0	39,3	2
Poplar	47,1	6,6	27,9	21,5	1
Pine	23,6	4,5	20,4	16,2	1
Spruce	19,3	8,5	11,8	29,7	1
Total	**11,9**	**5,1**	**11,5**	**20,9**	**100**

Because the aim of the research was to find a meaningful and efficient way to update an existing tree register, location accuracy was not contemplated. However, the same localizing method was used in the Seurasaari area in a separate study, where the standard error of the location was found to be about 1.3 meters.

4. Discussion

LiDAR data is acquired in Helsinki mainly for other urban mapping and measuring purposes. Urban tree mapping, therefore, provides a new way to utilize data that is already collected.

Tree mapping using ITD is the most critical phase of the updating process of a manual tree register. If a tree falls inside an inferior segment (a segment that is ill-shaped or representing a wrong tree), its location and both DBH and height might be affected. Inferior segments might represent only a small part of a single tree crown, or, in the worst case, the segment might contain parts of multiple trees or other urban objects such as cables, lamp posts etc. In such cases local maxima will not necessarily represent treetops of specific trees. LiDAR returns from incorrect trees or other objects result in poor quality LiDAR metrics and thus affect all predicted variables, including tree location. The variation in tree size is also a challenge in the segmentation process. Old and sparse broadleaved trees tend to have forked crowns. CHM resulted from these kind of crowns should be smoothed quite heavily in order to produce proper segments. The smallest trees, however, disappear into the background if CHM is smoothed too much. The disappearance occurs due to the rather narrow crowns of young trees, especially in leaf-off LiDAR data.

The accuracy of DBH estimates was good. In urban environment, Shrestha and Wynne (2012) reported an overall RMSE of 11,2 cm when LiDAR height metrics where used as explanatory variables in linear models. In Yu et.al. (2011) the RMSE of DBH was 21,49 % when Random Forest methods were used in forest environment. Accurate estimates may partly be a result of the homogenous nature of most roadside trees: there are long roadside sections consisting of very similar trees. However, due to the nature of Random Forest methodology, error of the DBH estimate for a single exceptional tree may be significant.

In an urban environment, the traditional tree characteristics such as volume and diameter lose some of their importance to attributes like tree location. In this study, tree location was determined by the XY-location of the highest LiDAR return inside the segment. In most cases, this method gives an accurate estimate of the tree top location (figure 1). However, the method may be misled by several things. Firstly, the location of the tree stem is rarely the same as the location of the tree top: urban deciduous trees are often unequally branched and their stems are often at an angle. Secondly, as mentioned earlier, the method is heavily influenced by the success of segmentation: LiDAR returns from nearby trees or other objects may cause considerable lack of precision in tree location. In future research, other ways of determining tree location should be studied.

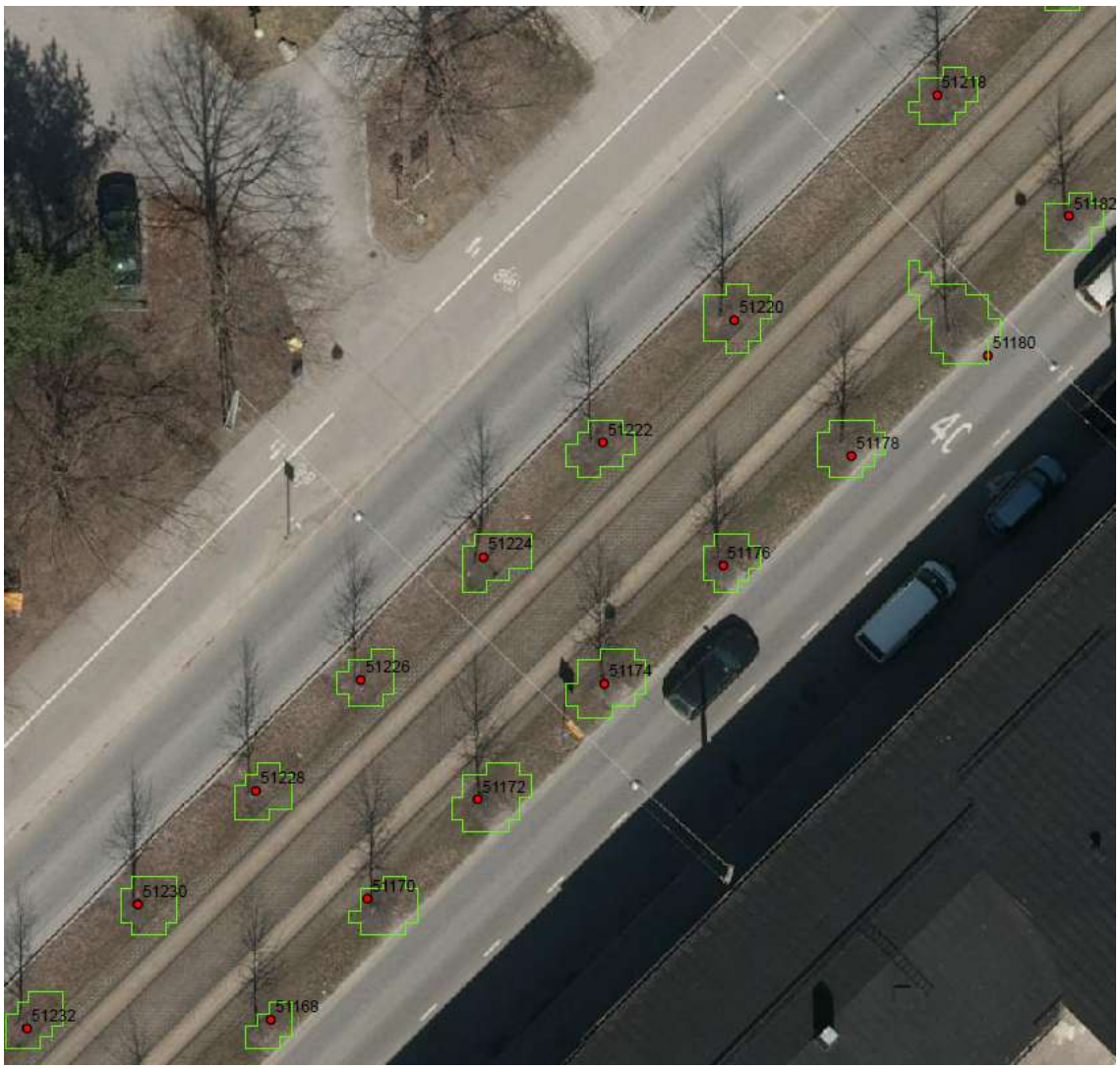


Figure 1: Segmented roadside trees. Red dots indicate the LiDAR based locations.

The location of the tree is the single most significant characteristic of a roadside tree, but other attributes are also worth observing. Before this study, the tree register of the City of Helsinki did not have information about tree height, as it has previously been too costly to determine for all trees. The LiDAR-based method created in this study allows cost-efficient determination of tree height among other LiDAR features. In an urban environment, tree height information can be utilized, for example, in traffic-oriented visibility analysis.

Although the stem location might differ from the treetop location, it is very useful for future updating. When applied to a new LiDAR dataset, treetops are most likely found in the vicinity of previously determined treetop locations. LiDAR accuracy in XY-plane is better than ~20 centimetres. When the treetop locations and their accuracy are known, joining register trees to new segments should be easier and only a small proportion of segments has to be manually combined to register trees. This makes LiDAR-based approach effective for maintaining tree registers. Once a segment is created and joined to a register tree, the tree should be found within the segment in future LiDAR data as well. This significantly enhances the productivity of the approach as the segmentation phase has to be executed only in areas that have newly planted trees. Also, the time-consuming joining of ill-located register trees and segments has to be done only once.

The quality of the tree register greatly affects the process. Joining of existing register trees and LiDAR segments is based on spatial location. The more the location information of a register tree differs from the reality, the more difficult it becomes to join the right register tree to the right segment. As mentioned earlier, the so-called "ghost trees" were also problematic. This was because the difference between old and present tree locations was usually very small (<1 m). As a result, multiple register trees ended up in the same segments, thus affecting the efficiency of combining register trees to segments.

When every register tree has its own segment, change detection can be carried out on a single-tree level. This enables effective monitoring and managing of urban trees. Single-tree biomass functions, for example, can be utilized in order to efficiently manage the carbon storage that the urban trees form. Traditional managing procedures benefit from exact registers as well: costly nursing actions can be allocated precisely, which reduces the costs of maintaining urban forested areas.

5. Conclusions

The LiDAR-based approach introduced above provides an efficient tool for updating urban tree registers. In a city environment, LiDAR data is widely used for other purposes which lower the costs of data acquisition. However, there are several elements that need to be improved. The definition of tree location, for example, needs further research. A combination of airborne LiDAR and Mobile Laser Scanning data would be most likely to improve the location accuracy.

Acknowledgements

Authors want to thank The City of Helsinki for materials and funding that made this study possible.

References

Bolund, P., Hunhammar, S. 1999. Ecosystem services in urban areas. Ecological Economics 29: 293-301.

Breiman, L. 2001. Random Forests. Machine Learning 45 (1): 5–32. Crookston & Finley 2007

Falkowski, M., Smith, A., Gessler, P., Hudak, A., Vierling, L. & Evans, J. 2008. The influence of the conifer forest canopy cover on the accuracy of two individual tree detection algorithms using lidar data. Canadian Journal of Remote Sensing 34 (2): 1–13.

Holmgren, J., Persson, Å. & Söderman, U. 2008. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. International Journal of Remote Sensing (29): 1537-1552.

Holopainen, M., Talvitie, M., 2006. Effects of Data Acquisition Accuracy on Timing of Stand Harvests and Expected Net Present Value. Silva Fennica 40: 531-543.

Hyyppä, J. & Inkinen, M. 1999. Detecting and estimating attributes for single trees using laser scanner. The Photogrammetric Journal of Finland 16: 27–42.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X. and M. Maltamo, 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. International Journal of Remote Sensing, 29(5), 1339-1366.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, BM & Wu, J. 2012. Remote Sensing 4: 950-974.

Koch, B., Heyder, U. & Weinacker, H. 2006. Detection of individual tree crowns in airborne lidar data. Photogrammetric Engineering and Remote Sensing 72: 357–363.

Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P. & Tokola, T. 2009. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. Silva Fennica 43(3): 507–521.

McRoberts, R.E., Cohen, W.B., Næsset, E., Stehman, S.V., Tomppo, E.O. 2010. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. Scandinavian Journal of Forest Research 254: 340-367.

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80: 88–99.

Næsset, E., 2004. Practical Large-scale forest stand inventory using a small-footprint airborne scanning laser. Scandinavian Journal of Forest Research 19: 164–179.

Omasa, K., Hosoi, T., Uenishi, T. M., Shimitsu, Y., Akiyama, Y. 2008. Three-Dimensional Modeling of an Urban Park and Trees by Combined Airborne and Portable On-Ground Scanning LiDAR Remote Sensing. Environmental Modeling and Assesment 13(4): 473–481.

Palenichka R. M., Zaremba, M. B. 2007. Multiscale Isotropic Matched Filtering for Individual Tree Detection in LiDAR Images. IEEE Transactions on Geoscience and Remote Sensing 45(12): 3944–3956.

Popescu, S., Wynne, R. & Scrivani, J. 2004. Fusion of small-footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, U.S.A. Forest Science 50: 551–565.

Shrestha, R., Wynne, R. H. 2012. Estimating Biophysical Parameters of Individual Trees in an Urban Environment Using Small Footprint Discrete-Return Imaging Lidar. Remote Sensing 4(2): 484–508.

Uuttera, J., Anttila, P., Suvanto, A., Maltamo, M. 2006. Yksityismetsien metsävaratiedon keruuseen soveltuvilla kaukokartoitusmenetelmillä estimoitujen puustotunnusten luotettavuus. Metsätieteen aikakauskirja 4/2006: 507–519.

Vastaranta, M., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H. & Viitala, R. 2010b. Determination of stands first thinning maturity using airborne laser scanning. Silvilaser 2010 conference proceedings.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T. and Maltamo, M. 2012. Comparative testing of single-tree detection algorithms under different types of forest. Forestry 85(1): 27–40.

Vauhkonen, J., Korpela, I., Maltamo, M., Tokola, T., 2010. Imputation of single-tree attributes using airborne laser scanning based height, intensity, and alpha shape metrics. Remote Sensing of Environment 114(6): 1263–1276.

Yu, X., Hyyppä, J., Vastaranta, M. & Holopainen, M. 2011. Predicting individual tree attributes from airborne laser point clouds based on random forest technique. ISPRS Journal of Photogrammetry and Remote Sensing 66: 28-37.

A Sensitivity analysis for a novel individual tree segmentation algorithm using 3D lidar point cloud data

Wei Yao[1], Peter Krzystek[1] & Marco Heurich[2]

[1]Department of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany, {Yao, Krzystek}@hm.edu
[2]Bavarian Forest National Park, 94481 Grafenau, Germany, marco.heurich@npv-bw.bayern.de

Paper Number: ### SL2012-106

Abstract

LiDAR sampling or full-area coverage is deemed as favorable means to achieve timely and robust characterizations of vertically distributed forest attributes. So far, two main strategies on the use of LiDAR data in forestry are reported: area-based method (ABA) and individual tree method (ITC). Recently, a novel 3D segmentation approach has been developed for extracting single trees from LIDAR data. It is an integrated approach, which delineates tree crowns by using the watershed algorithm and stem detection followed by applying 3D normalized cuts segmentation. However, all the parameters for the modules used in the whole strategy including watershed and normalized cut segmentations are empirically determined and the key parameter settings are to be optimized. Additionally, the robustness of the 3D tree segmentation approach needs to be examined in order to give a better understanding of the algorithm mechanism.

This paper highlights a study for sensitivity analysis of 3D single tree detection from small-footprint airborne LiDAR data by varying key parameters used in the segmentation procedure, such as segmentation threshold, minimal number of voxels allowed in a segment, voxel size, merged size of watershed segments, etc. The aim of the study is to find out the optimal combinations of key parameter values towards the performance of single tree detection via sensitivity analysis. It could help us to gain more insights into different models used in the whole procedure and is also of benefit to the further improvement and development in such models. Test data used in this work were captured with the Riegl LMS-Q680i scanner at four-fold point densities of 5pts/m^2, 10pts/m^2 15pts/m^2 and 20pts/m^2 under leaf-off condition. The study results proved the robustness and efficiency of the 3D segmentation approach in different airborne LiDAR data with respect to the single tree detection. Datasets whose point density is larger than 10 pts/m^2 seem to hardly much contribute to the improvement to the performance of 3D tree detection. The performance of the approach might be further revealed and improved by optimizing the determination of key parameters' values with respect to different data properties.

1. Introduction

Airborne laser scanning (ALS) or LiDAR has been widely used in mapping the Earth's surface and especially for forestry applications. Recent advances in LIDAR technology have demonstrated that new full waveform scanners can provide a higher spatial point density and additional information about the reflectional characteristics and vertical structure of trees (Stilla *et al*., 2007; Reitberger *et al*., 2008; Yao and Stilla, 2010 and Yao *et al*., 2010). As LiDAR instrument technology continues to improve and acquisition costs decrease, the wall-to-wall

characterization of large forested areas will become more common. Therefore, LiDAR sampling or full-area coverage is deemed as favorable means to achieve timely and robust large-area characterizations of vertically and horizontal distributed forest structures.

Two main strategies on the use of LiDAR data in forestry are reported: area-based method (ABA) and individual tree method (ITC). In the ABA, ALS point cloud data are aggregated at the inventory plot level by describing them based on the height distribution and canopy density metrics (Næsset, 2002), such as echo ratio, height percentiles and canopy cover percentiles. These metrics can be imported into regression models as independent variables. In this way, the response is an aggregation of single tree measurements on sample plots and could be as in the case of mean tree height, stem volume, biomass or forest fuel parameters (Nelson *et al*., 1988, Andersen *et al*., 2005 and Heurich and Thoma, 2008). In the ITC approach, either canopy height model (CHM) (e.g. Persson *et al*.,2002;Yu *et al*., 2011) or the ALS raw data points clouds (Reitberger *et al*., 2009; Vauhkonen et al., 2012; Yao *et al*., 2012) are used to be segmented into single tree objects. Properties at single tree level can subsequently be estimated using the tree segments, such as tree height, crown diameter, tree species, crown base height and stem volume. The ITC method is more intuitive than ABA since the derived parameters refer to the single tree, which is the smallest unit in forest management.

Novel methods for ITC approach tackle conceptually the segmentation problem with a 3D approach instead of using only the CHM. In combination with full waveform data Reitberger *et al*. (2009) successfully demonstrated that the overall detection rate of single trees could be significantly improved, especially in heterogeneous forest types. The improvement happened mostly to the lower forest layers with 20%. The fusion of 3D techniques with full waveform data pushed the single tree approach to a new level of completeness. Meanwhile, the estimation of tree shape parameters benefits from 3D tree segments. Moreover, the tree reflectional characteristics gain more insight into tree structure which are significant for tree species classification and growth condition identification. Therefore, unbiased species-specific models can be used to accurately determine forest biophysical parameters even at single tree level (Korpela et al., 2010 and Yao *et al*., 2012).

So far, little attention has been paid to the sensitivity analysis of detecting 3D individual trees from heterogeneous forest areas using LiDAR data. A fully automated method for tree detection is always desirable; a sensitivity analysis can be performed to characterize and quantify the uncertainty sources of the algorithm and enables to help optimize the selection of control parameters towards the best performance. Therefore, a detailed understanding of the uncertainty of the key control parameters on the performance of individual tree detection is required. In our former works of Reitberger *et al*. (2009), the smoothing factor and segmentation threshold were already investigated in a preliminary configuration with limited point density and value domain. In the past several years there are few authors dealing with the similar topics. In Gonzalez *et al*. (2010) and Lu *et al*. (2012), an uncertainty analysis is performed for estimating forest carbon stock and biomass from satellite imagery and LiDAR data. The dominate sources of uncertainty are the variation of input sample plot data and data saturation problem related to optical sensors. Monnet *et al*. (2010) conducted a sensitivity analysis of a treetop detection algorithm in the French Alps by automated evaluation of detection performance for several parameter combinations. The optimization of parameters may depend on both the laser data, mainly point density, and on the forest structure and species. Palleja *et al*. (2010) investigated the sensitivity of the tree volume estimates relative to different error sources in the spatial trajectory of a mobile terrestrial LIDAR and found out that the volume estimation is very sensitive to both errors in the determination of the distance from the LIDAR to the center of the trees and in the determination of the orientation angle of LIDAR. In Mutlu et al. (2008), fuel model maps obtained based on QuickBird image and LIDAR-derived variables were used to perform a

sensitivity analysis of fire behavior modeling. The study showed that the accuracy of fuel mapping can be improved by at least 13%.

The full waveform LiDAR systems have the possibility to overcome drawbacks of conventional laser scanners since they detect significantly more reflections for trees in the lower forest layer, and provide the intensity and the echo width as reflectional properties. The objectives of this work are (i) examine the influence of control parameters's uncertainty on the performance of tree detection in a mixed strand forest from airborne full-waveform LiDAR, (ii) exploit the feasibility towards optimizing control parameters for 3D single tree detection, and (iii) provide a preliminary result on sensitivity analysis of 3D tree detection in airborne LiDAR data.

The paper is divided into five sections. Section 2 focuses on the methodology of 3D segmentation of single trees and introduces key control parameters to undergo sensitivity analysis. Section 3 presents the experimental results which are obtained from full-waveform LiDAR data of four-fold point densities in Austrian floodplain forest. Finally, the results are discussed and concluded in the last two sections.

2. Method

2.1 Waveform decomposition

We assume that full waveform LIDAR data have been captured in a region of interest (ROI). A single waveform is decomposed by fitting mixed Gaussian pulse models to the waveform which contains N_R reflections (Figure 1).

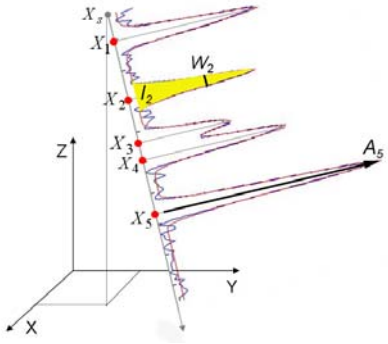


Figure 1: 3D points and attributes derived from a waveform (Reitberger *et al.*, 2009)

The vector $\boldsymbol{X}_i^T = (x_i, y_i, z_i, W_i, I_i)(i = 1,...,N_R)$ is provided for each reflection i with (x_i, y_i, z_i) as the 3D coordinates of the echo. Additionally, the points $\boldsymbol{X}_i$ are given the width $W_i = 2 \cdot \sigma_i$ and the intensity $I_i = \sqrt{2 \cdot \pi} \cdot \sigma_i \cdot A_i$ of the echo pulse with σ_i as the standard deviation and A_i as the amplitude of the reflection i (Reitberger *et al.*, 2008). Note that basically each reflection in the travel path of the laser beam can be detected by the waveform decomposition. This is remarkable since some conventional first/last pulse LIDAR systems have a dead zone of about 3 m which makes these systems effectively blind after triggering a reflection.

The LiDAR data are corrected by referencing W_i and I_i to the pulse width W^e and the intensity I^e of the emitted Gaussian pulse and correcting the intensity with respect to the run length s_i of the laser beam and a nominal distance s_0.

$$W_i^c = W_i / W^e \quad (1)$$

$$I_i^c = (I_i \cdot s_i^k) / (I^e \cdot s_0^k) \quad (2)$$

2.2 Single tree detection

The hybrid approach for single tree detection consists of two steps (Reitberger et al., 2009): watershed transformation with stem detection and normalized cut segmentation within merged watershed segments.

The coarse detection of single trees is achieved from the CHM by a watershed transformation. The CHM is derived by subdividing the ROI into a grid having a cell spacing of *cp* and N_C cells. Within each grid cell, the highest 3D point is selected and adapted with respect to the ground level z_j^{ground}, i.e. $z_j^{CHM} = z_j - z_j^{ground} (j = 1,...,N_C)$. The ground level z_j^{ground} is estimated from a given digital terrain model (DTM) by bilinear interpolation. In the next step, all the highest 3D points $\boldsymbol{X}_j^T = (x_j, y_j, z_j^{CHM})(j = 1,...,N_C)$ of all N_C cells are robustly interpolated in a grid that has N_X and N_Y grid lines and a grid size g_w. For this purpose an algorithm called 'gridfit' (D'Errico, 2006) is adopted which smoothens the surface by maintaining the surface gradients as small as possible. The trade-off between interpolation and regularization is determined by the adjustable smoothing factor. Both steps are carried out simultaneously in a least squares adjustment. The result is a smoothed CHM having equally spaced cells. The watershed segments derived on this CHM act as candidate regions where single trees could be contained. The results can also be improved by an additional stem detection method to further detect small trees which are not represented by local maximums.

Within every watershed segments the 3D segmentation technique using normalized cuts (Shi and Malik, 2000) is used to detect point clouds associated to single trees (Figure 2). This makes it possible to detect also smaller understory trees which cannot be indicated by local maxima in the CHM. This segmentation uses the positions (x_i, y_i, z_i) of the laser reflections and the pulse width W_i and the intensity I_i of the waveform decomposition. Additionally, stem positions or local maximums derived by the watershed segmentation of CHM are used as a-prior knowledge. The normalized cut segmentation applied to the voxel structure of a (merged) watershed segment is based on a graph *G* describing the adjacent topology between each voxel. The two disjoint segments *A* and *B* of the graph are found by minimizing the cost function:

$$NCut(A,B) = \frac{Cut(A,B)}{Assoc(A,V)} + \frac{Cut(A,B)}{Assoc(B,V)} \tag{3}$$

with $Cut(A,B) = \sum_{i \varepsilon A, j \varepsilon B} w_{ij}$ as the total sum of weights between the segments *A* and *B* and $Assoc(A,V) = \sum_{i \varepsilon A, j \varepsilon V} w_{ij}$ as the sum of the weights of all edges ending in the segment *A*. The weights w_{ij} specify the similarity between the voxels and are a function of the LiDAR point distribution and various features. The segmentation is realized by maximizing the similarity within the segment voxels and minimizing the similarity between the segments *A* and *B*. It turned out that the spatial distribution of the LiDAR points mainly influences the weighting function. The features derived from the LiDAR points attributes of W_i and I_i only support in second instance the segmentation result. Note that the 3D segmentation approach is not limited to full waveform LiDAR data. It can also successfully be applied to conventional LiDAR data just providing 3D point coordinates.

2.3 Key control parameters for sensitivity analysis

Several key control parameters for 3D tree extraction algorithm from airborne LiDAR data are selected here to perform the sensitivity analysis.

2.3.1 Threshold for normalized cut value $NCut_{\text{Thres}}$

The minimization of *NCut(A,B)* is solved by the corresponding generalized eigenvalue problem $(D-W)y=\lambda Dy$, where the minimum solution y_1 corresponds to the second smallest eigenvalue λ_1. Since y_1 is real-valued, but may only have two distinct indicator values (+1,-1) we need to binarize it by introducing a threshold into the histogram of y_1. Using simply the values 0 or *median*(y_1) leads to proper results. The results can be improved by testing several possible values with respect to the resulting value of *NCut* and choosing the value causing the smallest *NCut* value. We find the optimal value by stepwise testing all values within the range of y_1 in a small interval. Thus, the graph *G* is subdivided into two disjoint segments $G_1$and G_2. After that, the abovementioned steps of solving the eigenvalue problem and binarization of y_1 are to be performed iteratively on the sub-graphs $G_1$and G_2 until the value for *NCut* reaches or exceeds the threshold $NCut_{\text{Thres}}$.

2.3.2 Minimal number of voxels in a segment $N_{\text{Seg\_Min}}$

Likewise, the control parameter N_{Seg_Min} is also used to terminate the iteration of the subdivision of the graph *G*, if the number of voxels in one of divided sub-graphs is smaller than the certain minimal number N_{Seg_Min}.

2.3.3 Voxel size V_{Size}

To enable a 3D tree segmentation in LiDAR point clouds based on normalized cuts the ROI delineated by the watershed segmentation is subdivided into a voxel structure with a voxel spacing of V_{Size} and $N_v = N_v^x \times N_v^y \times N_v^z$ voxels (Figure 2). Within each voxel of size V_{Size} we collect *N* reflections $X_i^T = (x_i, y_i, z_i, W_i, I_i)$ $(i = 1....N)$, where only voxels comprising at least one reflection are used in the segmentation. The voxel structure is represented as a region adjacency graph $G=\{V,E\}$ with *V* as the voxels representing the nodes and *E* as the edges formed between every pair of nodes. The similarity between two nodes $\{i,j\} \in V$ is described by the weights w_{ij} which are computed from features associated with the voxels. Basically, the similarity between voxels decreases with increasing distance between two voxels and drops down to zero beyond the threshold r_{XY} for the mutual distance in order to keep the graph *G* at a reasonable size for computation. The goal of normalized cut segmentation is to divide the graph *G* into disjoint voxel segments *A* and *B* (Figure 2) by maximizing the similarity within the segment members and minimizing the similarity between the segments *A* and *B*.

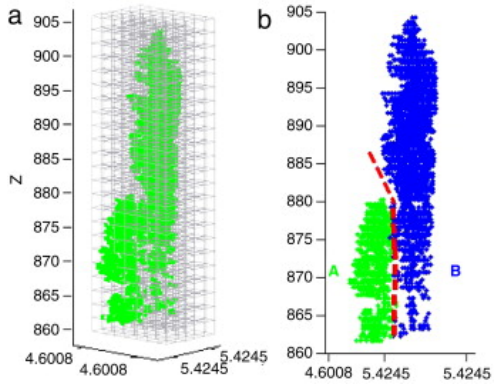


Figure 2: (a) Subdivision of the ROI into a voxel structure and (b) division into two tree segments. (Reitberger *et al*., 2009a)

2.3.4 Merged size of watershed segments S_{Seg_Merg}

The watershed transformation is performed at the first step to segment the CHM and delineate 2D crowns of dominate trees. The watershed segments can be viewed as candidate areas for 3D tree segmentation under the CHM. To overcome the oversegmentation and maintain the computational efficiency of the algorithm on the large-area forest, the adjacent watershed segments are usually merged to build a larger area to undergo the normalized cuts segmentation in 3D. The parameter S_{Seg_Merg} controls the maximal size of the merged watershed segments. The merging process of adjacent watershed segments will be terminated until the size of merged segments reaches this threshold. It is expected to see which kind of impact on the performance of 3D tree detection the merged sized could have. An illustrative example for merging watershed segments can be found in Figure 3.

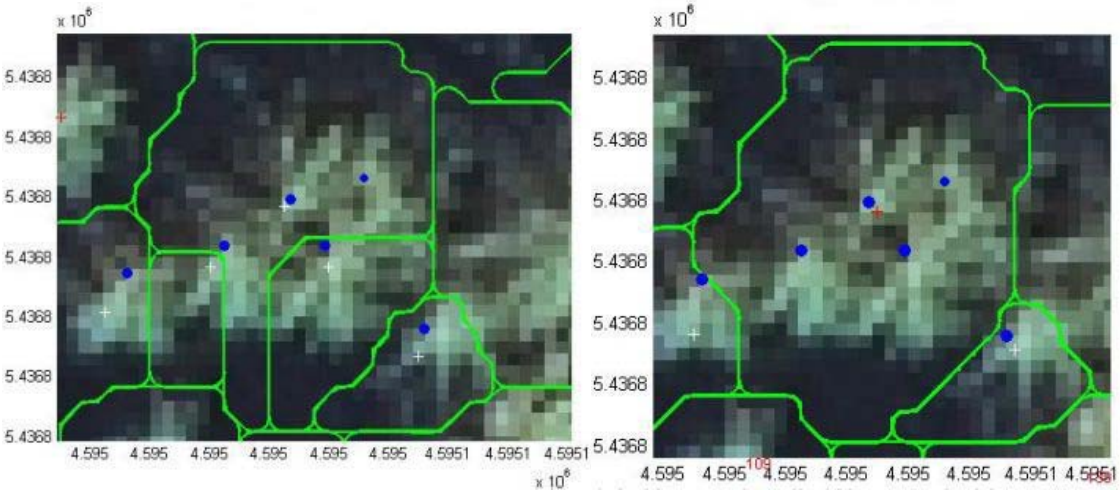


Figure 3: Left: original watershed segments; Right: merged watershed segments.

3. Experiment

3.1 Material

Experiments were conducted in the Austrian floodplain forest, which is located in western Austria along the border to the Germany (47° 58' 19" N, 12° 55' 9" E). Twenty seven sample plots with an area size between 200 m^2 and 500 m^2t. The forest comprises alluvial forest, spruce monocropping, deciduous mixed forest with partial pinewood. Practically, all age classes are to be found. Furthermore, the trees are subdivided into three layers with respect to the top height h_{top} of the plot, where h_{top} is defined as the average height of the 100 highest trees per ha. The lower layer contains all trees below 50% of h_{top}, the intermediate layer refers to all trees between 50% and 80% of h_{top}, and finally, the upper layer contains the rest of the trees. The plots comprise forest in the early and the late pole phase. Table 1 summarizes the characteristics of the individual sample plots.

Table 1: Characteristics of sample plots

Plot	Size [ha]	Altitude [m]	Trees/ha	Deciduous [%]	N lower layer	N interm. layer	N upper layer
I	0.05	441	448	0	0	0	12
II	0.04	441	483	0	0	0	9
III	0.05	441	417	100	0	2	11
IV	0.04	441	349	100	0	1	7
V	0.05	441	490	0	0	0	13
VI	0.04	440	261	100	0	2	2
VII	0.04	440	202	100	0	1	4
VIII	0.03	441	560	75	0	6	2
IX	0.05	441	453	0	0	0	14
X	0.05	440	441	0	0	0	13
XI	0.04	440	272	100	0	0	5
XII	0.05	439	196	100	0	0	6
XIII	0.05	439	487	0	0	0	14
XIV	0.05	439	490	0	0	0	13
XV	0.06	439	679	0	0	0	23
XVI	0.04	440	371	100	0	2	7
XVII	0.05	439	698	0	0	0	18
XVIII	0.05	482	576	0	0	0	16
XIX	0.05	483	633	12	0	0	6
XX	0.05	468	631	94	9	4	3
XXI	0.05	464	405	90	4	2	3
XXII	0.05	464	690	0	0	0	17
XXIII	0.04	448	330	0	0	0	5
XXIV	0.05	447	471	83	1	2	6
XXV	0.05	456	692	0	0	0	19
XXVI	0.04	442	340	0	0	2	6
XXVII	0.04	442	411	0	0	1	7

Full waveform data have been collected by Forest Mapping Management GmbH in Salzburg with the Riegl LMS-Q680i in March 2011 after snowmelt but prior to foliation with four-fold average point density of 5-20 points/m^2 (Table 2). The laser pulse repetition rate was 240 kHz, the pulse width at half maximum reached 4 ns and the laser wavelength was 1550 nm. The flying altitude of 700 m resulted in a footprint size of ca.30 cm.

Table 2: Configurations of airborne LiDAR campaign

Time of flight	March 2011
Foliage	Leaf-off
Scanner	Riegl LMS-Q680i
Point density: Pts/m^2	5,10,15,20
AGL [m]	700
Beam divergence(mrad)	<= 0.5
Scan angle	22.5°

3.2 Calibration

Unfortunately, there were no special calibration flights performed in this data acquisition campaign to determine the calibration of the Riegl full waveform system, where several tracks were flown at different flying heights along and across the airfield. Therefore, in this work the mean intensity I_i was calculated by the correction only with respect to the emitted intensity I_e, and the mean run length s_i for each track according to Eq. (2), assuming simply the coefficient k = 2 for all flights.

3.3 Field data and evaluation

We applied the strategy to the sample plots acquired under leaf-off condition and validate the results by the method of single-tree based evaluation. Reference data for all 327 trees with DBH larger than 15 cm were collected for 218 Norway spruces (Picea abies), 14 Scots pines (Pinus sylvestris), 12 Canadian poplars (Populus×Canadensis), 26 Oaks (Quercus), 2 European beeches (Fagus sylvatica), 22 European ashes (Fraxinus excelsior), 5 Black Alders (Alnus glutinosa), 27 Sycamore maples (Acer pseudoplatanus), and 1 lime tree (Tilia Europaea). Tree parameters including the DBH, total tree height, stem position and tree species were measured and determined by GPS, tachometry and the 'Vertex III hypsometer' system. Additionally, a DTM with a grid size of 1 m and an absolute accuracy of 25 cm was available for the test sites.

The single tree based evaluation is performed based on single-tree level by finding matched tree segments in the reference. The tree detection results are evaluated by comparison with reference data using two criterions: i) the distance of detected trees should be smaller than 60% of the mean tree spacing of the plot; ii) the height difference between detected and reference trees should be smaller than 15% of h_{top}. If a reference tree is assigned to more than one tree position, the tree position with the minimum distance to the reference is selected. Detected trees that are liked to one tree position are so-called "true positives" and detected trees without any link to a tree position are treated as "false positives". We used the measures Completeness and Correctness, to characterize the performance of the individual tree detection, which are defined in equations (4 and (5), respectively. The descriptive statistics of the field trees are summarized in Table 3.

$$Completeness = \frac{Number_{true\ positives}}{Number_{reference\ trees}} \tag{4}$$

$$Correctness = \frac{Number_{true\ positives}}{Number_{detected\ trees}} \tag{5}$$

Table 3: Field data for single tree based evaluation

Time of acquisition	Mar. 2011	
	Tree height(m)	DBH(cm)
Min	1.0	2.0
Max	44.0	50.0
Mean	21.2	22.8
Standard deviation	8.7	10.1

3.4 Results

The procedures for 3D single tree detection were applied to the test plots in a batch procedure without any manual interaction. The experimental results describing the impact of each selected control parameter on the performance of tree detection are illustrated in Figures 4-7. Parameters that were maintained as constant values during the sensitivity analysis included the following: $NCut_{\text{Thre}} = 0.2$, $N_{Seg\_Min} = 16$, $V_{Size} = 0.5$, $S_{Seg\_Merg} = 1000$.

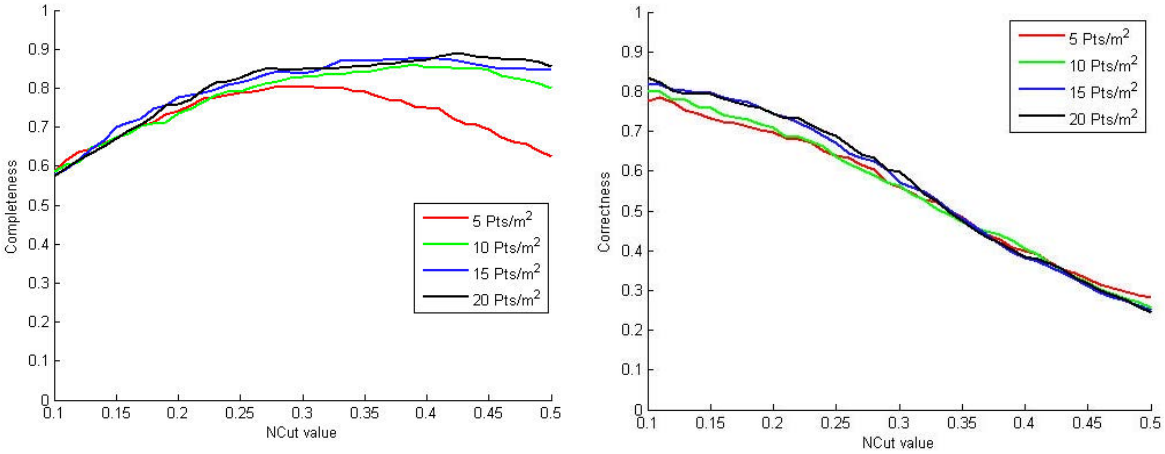


Figure 4: Sensitivity analysis with respect to the parameter $NCut_{\text{Thres}}$, left: completeness; right: correctness

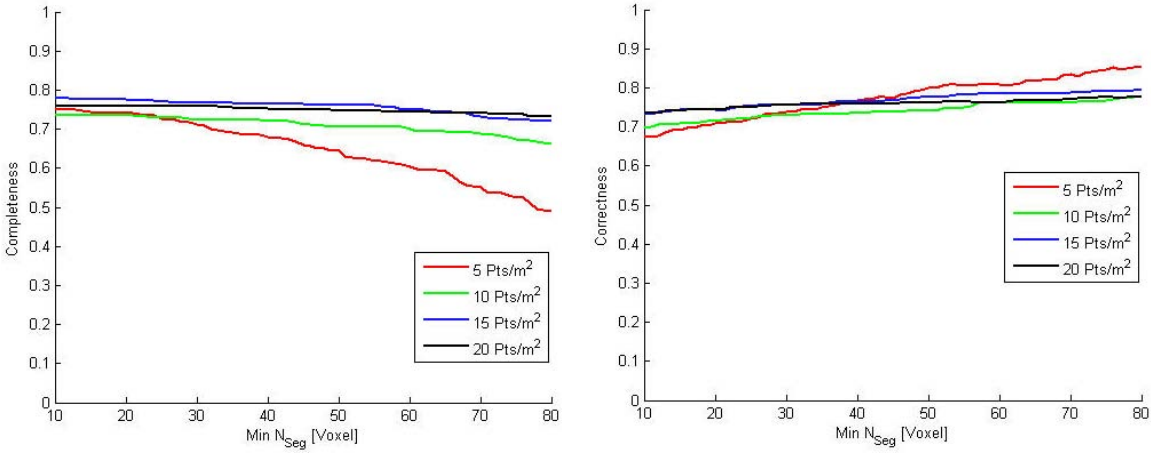


Figure 5: Sensitivity analysis with respect to the parameter N_{Seg_Min}, left: completeness; right: correctness

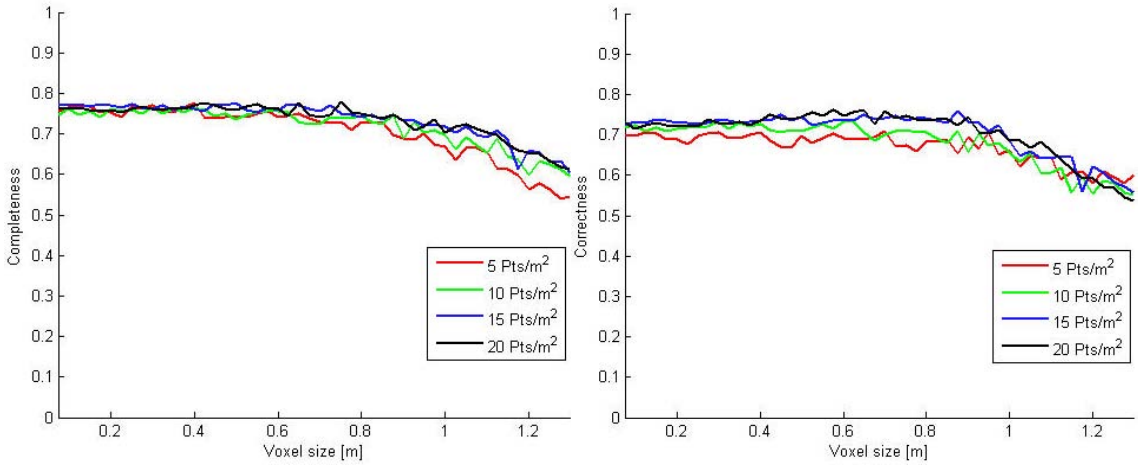


Figure 6: Sensitivity analysis with respect to the parameter V_{Size} , left: completeness; right: correctness

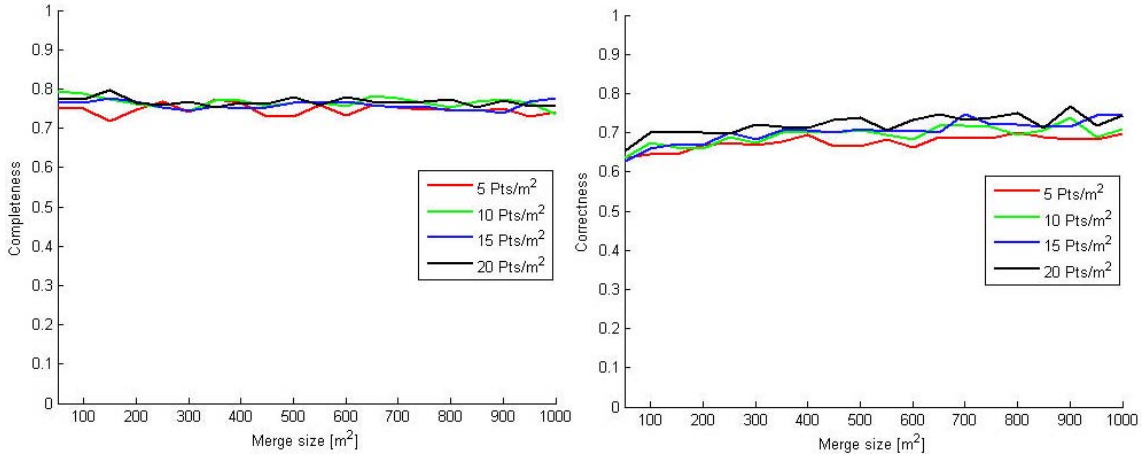


Figure 7: Sensitivity analysis with respect to the parameter S_{Seg_Merg} , left: completeness; right: correctness

4. Discussion

The experimental results of the sensitivity analysis of the new 3D tree detection strategy on airborne LiDAR data with four-fold point density demonstrate that the threshold for normalized cuts segmentation $NCut_{\text{Thres}}$ plays the most significant role in the determination of overall performance of 3D tree detection from airborne point clouds.

The larger $NCut_{\text{Thres}}$ values lead to more tree segments with a higher completeness but lower correctness (oversegmentation), since the $NCut_{\text{Thres}}$ value defines the threshold for the evaluating the dissimilarity measure between sub-graphs. It can also be seen that the low point density dataset such as one-fold (5pts/m^2) or two-fold (10pts/m^2) point density data could even produce a descending trend upon the completeness when $NCut_{\text{Thres}}$ value is larger than 0.4. Such low point density data set has reached their oversegmentation limit and will not produce any more new segments by using bipartition cuts. However, the performances of 3D tree segmentation for the data sets of four different point densities are almost identical if the $NCut_{\text{Thres}}$ value being smaller than 0.25 is selected. When viewing the results of impact of variation of N_{Seg_Min} on the performance of 3D tree detection, we attain the conclusion that the parameter N_{Seg_Min} has hardly

distinct influence on the results if the three-fold and four-fold point density LiDAR data are used to segment individual trees. Only a slight increased correctness is observed. On the other side, for both the one-fold and two-fold point density data, the completeness of the results has degraded gradually with respect to the increased N_{Seg_Min} value while the correctness has increased, which is actually an undersegmentation effect. The results from the variation of the voxel size V_{Size} seems to show that there is only a slight consistent impact of V_{Size} on the performance of the 3D tree detection leading to a slightly degraded performance when V_{Size} increases, since the progress of completeness and correctness curves follow a fluctuant distribution around the same value. The higher point density cannot guarantee a better performance or stable results for tree detection if the voxel size varies. Moreover, if the voxel size increased to be more than 0.8m an undersegmentation happened to the one/two-fold point density dataset. It could be due to the reason that the large voxel size reduced the dimension of the adjacency graph and some small trees have been unified into one graph node. For the parameter of merged size S_{Seg_Merg}, the experimental results of sensitivity analysis also showed that it has no significant impact on the results of the algorithm. The progress of performance curves (Figure 7) more or less looks like stochastic fluctuation around a constant value. The point density of the LiDAR data also seems not to play a role at this stage. In this case, it is difficult to state that whether the high point density can really be helpful to improving the performance of tree detection algorithm, at least for the data with a point density being larger than 10 pts/m^2. Overall, the optimal value for $NCut_{\text{Thres}}$ can be selected as between 0.16 and 0.23 in order to achieve the trade-off between completeness and correctness. Generally, datasets whose point density is larger than 10 pts/m^2 can hardly contribute to a significant improvement to the performance of 3D tree detection. Moreover, all other control parameters except $NCut_{\text{Thres}}$ are almost insensitive w.r.t. correctness and completeness if the point density is larger than 5 pts/m^2.

5. Conclusion and outlook

It is necessary to quantize the influence of uncertainty associated with various control parameters of the single tree detection based on normalized cuts, since the performance of the forest characterization is directly related to the accuracy of tree detection algorithm. The parameter setting of 3D normalized cuts segmentation for tree detection needs to be optimized towards different data properties and forest structures, such as point density, forest type and tree growth phase. The experimental results of this works showed that the threshold for normalized cut value $NCut_{\text{Thres}}$ is the most important control parameter in the whole algorithm which is most sensible to the performance of tree detection and needs to be carefully adjusted. Future works could be focussed on the simultaneous optimization of various control parameters and the extension to the sensitivity analysis for tree species classification and forest parameter estimation. Finally, for comparison purpose we will conduct a sensitivity analysis on LiDAR data acquired in other forest types with different stem densities.

Acknowledgement

The authors want to thank Mr. Ing. Hermann Novak from Forest Mapping Management GmbH in Salzburg, Austrian for providing us the experimental and field reference data.

References

Andersen, H. E., McGaughey, R. J. and Reutebuch, S. E., 2005. Estimating forest canopy fuel parameters using LIDAR data. Remote Sensing of Environment, 94(4), 441−449.

D'Errico, 2006 D'Errico, J., 2006, Surface fitting using gridfit. http://www.mathworks.com/matlabcentral/fileexchange(accessd on 10.01.2009)

Gonzalez, P., Asner, G.P., Battles, J., Lefsky, M.A., Waring, K M. and Palace, M., 2010. Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California, *Remote Sensing of Environment*, 114(7), 1561-1575.

Heurich, M. and F. Thoma., 2008. Estimation of forestry stand parameters using laser scanning data in temperate, structurally rich natural beech (Fagus sylvatica) and spruce (Picea abies) forests. *Forestry*, 81. 645-661.

Heurich, M., 2008. Automatic recognition and measurement of single trees based on data from airborne laser scanning over the richly structured natural forests of the Bavarian Forest National Park. *Forest Ecology and Management*. 255 (2008) 2416–2433.

Korpela, I., Ørka, H.O., Maltamo, M., Tokola T. and Hyyppä, J., 2010. Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, 44(2), 319–339.

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang M., Laurin, G V. and Saah, D., 2012. Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates. *International Journal of Forestry Research*, vol. 2012, Article ID 436537, 16 pages.

Monnet, J-M., Merminy, E., Chanussotz, J. and Berger, F., 2010. Tree top detection using local maxima filtering: a parameter sensitivity analysis. Silvilaser 2010, Freiburg, 14th-17th Sept. 2010.

Mutlu, M., Popescu, S.C. and Zhao, K., 2008. Sensitivity Analysis of Fire Behavior Modeling with LIDAR‐Derived Surface Fuel Maps. *Forest Ecology and Management*, 256, 289‐294.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80(1), 88-99.

Nelson, R., Krabill, W. and Tonelli, J., 1988. Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, 24(2), 247−267.

Palleja, T., Tresanchez, M., Teixido, M., Sanz, R., Rosell, J.R. and Palacin, J., (2010). Sensitivity of tree volume measurement to trajectory errors from a terrestrial LIDAR scanner. *Agricultural and Forest Meteorology*, 150(11), 1420-1427.

Persson, A., Holmgren, J. and Soderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, 68(9), 925−932.

Reitberger, J., Krzystek, P. and Stilla, U., 2008. Analysis of full waveform LiDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing*, 29(5), 1407 – 1431.

Reitberger, J., Schnörr, C., Krzystek, P. and Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), 561–574.

Shi, J. and Malik, J., 2000. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888 – 905.

Stilla, U., Yao, W. and Jutzi, B., 2007. Detection of weak laser pulses by full waveform stacking. In: Stilla U, et al (eds) PIA07 Photogrammetric Image Analysis 2007. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, Vol 36(3/W49A), 25-30.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T. and Maltamo, M., 2012. Comparative testing of single tree detection algorithms under different types of forest. *Forestry,* 85(3), 27-40.

Yao, W. and Stilla, U., 2010. Mutual Enhancement of Weak Laser Pulses for Point Cloud Enrichment Based on Full-Waveform Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 48(9), 3571-3579.

Yao, W., Krzystek, P. and Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment*, Vol.123, 368-380.

Yao, W., and Stilla, U., 2011. Comparison of two methods for vehicle extraction from airborne lidar data toward motion analysis. *IEEE Geoscience and Remote Sensing Letters* 8(4), 607–611.

Analysis of effects of scanning angle on ALS-derived vegetation metrics in a nationwide airborne ALS acquisition

Alessandro Montaghi, Mikael Egberth, Jörgen Wallerman,
Mats Nilsson, Johan Holmgren & Håkan Olsson

Section of Forest Remote Sensing, Department of Forest Resource Management,
Swedish University of Agricultural Sciences, Umeå, Sweden

alessandro.montaghi@gmail.com, Mikael.Egberth@slu.se, Hakan.Olsson@slu.se

Paper Number: SL2012-124

1. Introduction

In the summer of 2009, the Swedish government, under the co-ordination of Lantmäteriet (the Swedish mapping, cadastral and land registration authority), started a five-year project (2009-2013) using airborne laser scanning for production of the New National Elevation Model (in Swedish: Ny Nationell Höjdmodell, NNH) for all of Sweden (450,000 km^2). The primary aim of this project is to produce a 2 meter grid DEM (Digital Elevation Model) in which the standard error is better than 0.5 m. A total of about four hundred scanning areas, with a size of 25 km by 50 km, are being scanned with a nominal density of 0.5-1 point per square meter, and with a maximum scanning angle of ± 20 degrees (Petersen and Burman Rost, 2011). The acquisition is primarily done using Leica Geosystems ALS50-II and ALS60 sensors, with Optech ALTM Gemini as a complementary sensor. In each scanning area, 21 parallel flight lines were flown, with a nominal overlap of 20%, in addition also three perpendicular crossing lines where acquired. This flight line arrangement was designed to allow strip adjustment techniques, based on sensor parameter calibration, to be used for creation of a seamless final product (Toth, 2009). Finally, main and cross strips were merged together and delivered in blocks of 2.5 km by 2.5 km.

The ALS data acquired for the NNH project is a resource of interest also for forest estimation. However, since the ALS survey is being carried out for a purpose other than measuring forest parameters, there are a number of issues that need to be considered. These include the effect of different time points for the scanning, the relatively large view angles used, and the positional accuracy for the NFI plots. In this paper, the effect of one of these issues, the view angle influence on LiDAR forest metrics, will be reported. At the conference also initial results from predictions based on LiDAR data and SPOT HRG satellite data, trained with national forest inventory sample plots, will be presented.

2. Method

2.1 Study area

The study was conducted over two sites of Sweden (see Figure 1). One study site was located in the northern part of Sweden, in the province of Norrbotten where the study area is approximately 7,500 km^2. The central point of the analyzed area is located at Lat. N 65° 50', Lon. E 21° 43'. The topography ranged from 0 to 590 m above sea level (a.s.l.). The second site was situated in the southern part of Sweden (Lat. N 56° 27', Lon. E 13° 35', 0 – 290 m a.s.l.) with a surveyed area of approximately 8,162 km^2. The forests of the northern site are largely

dominated by coniferous forest, while the forests cover types vary more in southern Sweden, with a major presence of several deciduous species as well as coniferous.

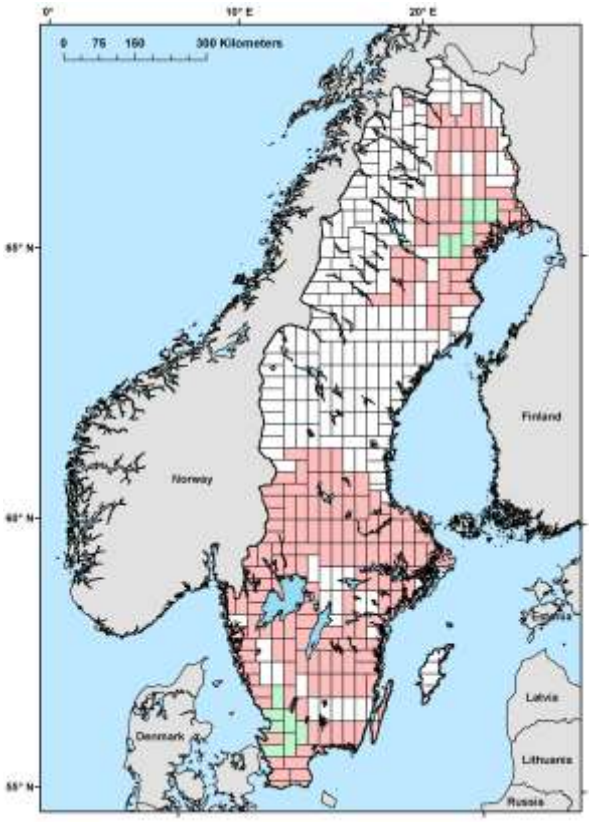


Figure 1: For the production of the New National Elevation Model of Sweden (NNH), the whole country was divided in approximately four hundred scanning areas (rectangles with black outline), each with a size of 50 km by 25 km, with scanning schedule according to various priorities. At the end of 2011, a total of 192 scanning areas were acquired (red rectangles), and from these data 13 areas (green rectangles) were selected for this study.

2.2 Airborne Laser Scanning data acquisitions and processing

From the NNH data, we selected six scanning areas for the northern site and seven scanning areas for the southern site. The data acquisition for the Northern study site was performed between August 13th 2009 and September 27th 2009. The ALS data were acquired by an ALS60 system for five areas, and by an ALS50-II system for the remaining scanning area. The southern site was acquired between April 2nd 2010 to July 1st 2010, using ALS50-II for four areas and ALS60 for three areas. Flying height varied between 1,740 m and 2,400 m, flight speed between 268 km/h and 287 km/h, and a pulse rate between 70 and 104 KHz (Petersen and Burman Rost, 2011).

The collected data were post-processed by the company Blom AS. A progressive Triangulated Irregular Network (TIN) densification method, developed by Axelsson (2000) and implemented in Terrascan software, was used to classify the point cloud into ground, water and above-ground returns. The laser hits finally classified as ground and water returns were used by Terrascan to generate the DEMs with a grid size of 2 by 2 m.

In this study, the ALS points classified above-ground were normalized through the subtraction of the laser hits' height with the pixels' DEM values at the corresponding locations. Subsequently, a fixed height threshold of 2 m was applied to eliminate ground pulses incorrectly classified as above-ground and to avoid the presence of boulders and low-lying vegetation (Nilsson, 1996). Additionally, an upper threshold of 40 m was applied to eliminate any noise from above the canopy.

A square grid with cells of 100 x 100 m was laid over each crossing between the main and crossing strips, and in each square a sample point was randomly inserted. Then, for each point a

square of 10 m by 10 m was created having a random orientation. Squared samples were intersected with the 1:100,000 topographic line map of Sweden and only those samples belonging entirely to the forest were considered for extraction of ALS metrics. These samples were used to select laser returns from the parallel flight lines and the perpendicular crossing lines, and since the above-ground returns were extracted for forested areas, they could be defined as Canopy height. Samples were rejected if the returns had scanning angles greater than zero degrees for the perpendicular crossing lines and if all returns did not have the same scanning angle value for the parallel flight line. Furthermore, samples were rejected if the data survey between perpendicular crossing lines and parallel flight lines was not carried out on the same day. This last condition allows us to assume that there were no changes in the forest condition inside the samples. In order to minimize the influence of young plantations, only plots with a height at the 99th percentile of $\geq$ 7 m were considered. Finally, these selections reduced the number of samples for the Northern (n = 1309) and Southern site (n = 1001), with each sample flown twice on the same day: one time always with a scanning angle of 0 degrees (cross strips) and a second time with an absolute scanning angle degree ranging from 0 degrees to $\geq$ 20 degrees (main strips), (see Figure 2).

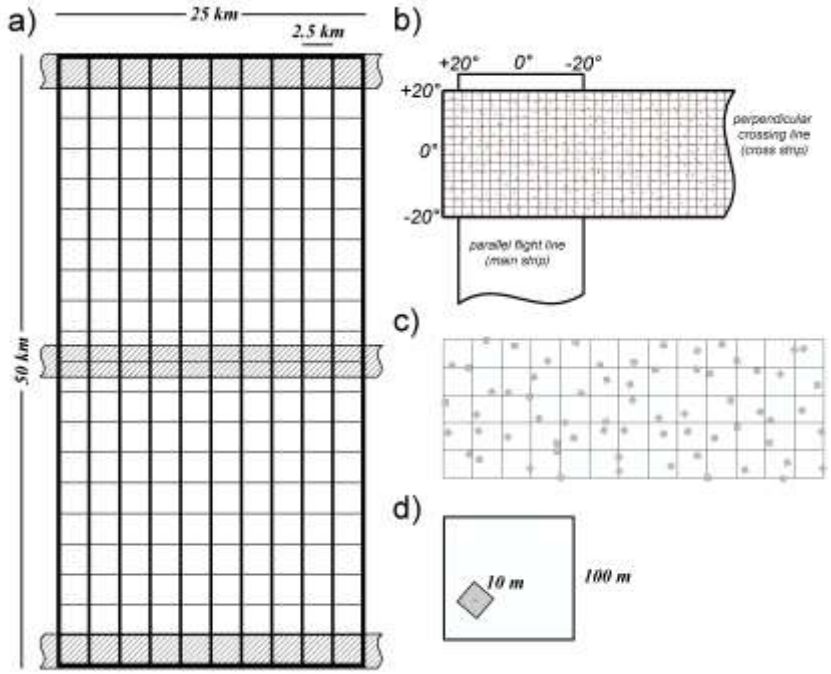


Figure 2: Design of the sampling scheme. Each scanning area of 25 km by 50 km (a) derived from the union of 21 parallel flight lines (main strips) and 3 three perpendicular crossing lines (cross strips), and the data are delivered in square blocks with 2.5 km to a side. In each crossing between main and cross strips, a square grid with cell of 100 x 100 m was overlain (b, c), and in each square, a sample point was randomly inserted (red dots). A square with 10 m to a side with random orientation (d) was created for each point in order to sample the laser return (note that objects in the Figure are not to scale).

Thirty-two plot-level LiDAR metrics were extracted from the Canopy height, both for the perpendicular crossing lines and for the parallel flight lines with a script coded in the R language (R Development Core Team, 2011). The metrics were selected according to previous studies reported in the literature (Evans et al., 2009; Hopkinson and Chasmer, 2009; Næsset, 2002, Parker and Russ, 2004). All returns were used to increase the sampling density of pulses, and metrics belonging to five categories were extracted. The categories used in this paper were *i) distributional metrics*, namely Mean (m) and Modal (m) height, Standard Deviation (m) of

height, Median Absolute Deviation (m) and Average Absolute Deviation (m) of height, Interquartile range of height (m), and Skewness and Kurtosis of height; *ii) height metrics* (in m), namely Maximum height and percentile of LiDAR height distribution (10^{th}, 25^{th}, 50^{th}, 75^{th}, 80^{th}, 90^{th}, 95^{th}, and 99^{th} percentiles); *iii) layer density metrics* (in %), namely Cumulative Density and Stratified Density of all returns; *iv) structural descriptor metrics* (in %), namely Understory ratio, Vegetation ratio, Canopy Relief ratio, and Point Height Diversity index; and *v)* other metrics computed for the aims of this study defined as *point density metrics* (in points per square meters), such as number of all returns, number of ground returns, number of all canopy returns, number of canopy returns above a given height of ≥ 2 m, and Understory returns (number of canopy pluses between 0 and < 2 m).

2.3 Statistical Analysis

Statistical analyses were performed using functions implemented in R. To explore a possible effect of forest structure and sensor type, the ALS metrics derived from the ALS60 and ALS50-II sensors were analyzed together for each study site because previous analysis showed no appreciable differences in function of the sensor types. The ALS60 model represents an upgrade of the ALS50-II, and the Leica scanners used in the national survey use very similar sensor technology (Vastaranta et al., 2011). Shapiro-Wilk's test of normality for all scanning angle classes was first used to test for normality in each metric. The test indicated that the majority of the metrics' variables violated this assumption, and as a result, a non-parametric Wilcoxon signed-rank test was employed to determine whether the single metrics were statistically significant inside each scanning angle class. The following hypotheses were tested for each ALS metric:

H_0: No difference exists between metrics derived from a scanning angle of 0 degrees and metrics derived from an absolute scan angle that ranges from 0 degrees to 20 degrees;

H_a: A difference exits between metrics derived from a scanning angle of 0 degrees and metrics derived from an absolute scan angle that ranges from 0 degrees to 20 degrees.

If they were significant ($p < 0.05$), then H_0 was rejected and the alternative hypothesis was accepted. Accepting the alternative hypothesis would indicate an influence of scanning angle on the repeatability of ALS metric to describe the forest structure.

3. Result

Results are shown in Tables 1 and 2 that were obtained using Wilcoxon signed-rank test for each category of metrics.

3.1 Distributional metrics

From an inspection of the results, it appears that the distributional metrics were generally stable (p = not significant) for all scanning angle configurations tested. However, significant differences ($p < 0.05$) were observed for all metrics inside the classes without an evident correlation with the scanning angles.

3.2 Height metrics

The plot-level maximum height and the height percentiles were relatively stable. Nonetheless, the tests were significantly different in some classes, especially for the height percentiles from the class 0°-15° to 0°-20° in the northern site, and in the classes 0°-12° and 0°-13° for the southern site.

3.3 Layer density metrics
A visual examination of the density metrics showed that the Cumulative Densities were more stable if compared with the Stratified Densities due to the fact that the former represented the cumulative number of pulses. However, the significant differences of the Stratified Densities were located in some of the classes where other categories of metrics were significantly different. This result explains in part the variance within the plot and is not an explicit influence of the scanning angle.

3.4 Structural descriptor metrics
The results of the Wilcoxon signed-rank test indicated that for the two sites, the significant differences showed a similar trend. In particular, the Understory and Vegetation ratio were strongly and significantly different ($p << 0.01$) for the class 0°-10°. The indices Canopy Relief ratio and Point Height Diversity were found to be stable among the classes, with the exception of some classes where the significant difference seems to be related to the internal variability of the forest plots used in the analysis.

3.5 Point density metrics
The ALS-metrics belonging to this category describe the density of the laser returns intercepted by the vegetation and by the ground. These metrics were stable only from class 0°-0° to 0°-5°, whereas after these classes the metrics were strongly and significantly different ($p << 0.01$). However, the ground returns showed a greater stability among the classes, potentially explaining the significant difference and the similar trend observed for the Understory and Vegetation ratios

4. Discussion

The results for both sites showed a scanning angle influence on the metrics belonging to the "Point density metrics" category and on the Understory and Vegetation ratios, while the remaining categories had significant differences for some distinct classes in the Northern and Southern sites.

The variations in the Understory and Vegetation ratios may be inherent to the differences in point densities between scanning angles of 0 degrees and ≥ 10 degrees, because all Point density metrics are affected by the mean increment of the sampling density, whereas ground returns are affected only to a minor degree with respect to metrics related to vegetation. The relative lower increment in ground point density explains the variation of the Understory and Vegetation ratios. The reason behind this lower increment in ground point density is unclear and may be inherent to the flight system setting and sensor type (i.e,. ALS50-II and ALS60 are used for corridor mapping ALS applications), which are optimized for ground point detection in order to map the ground elevation with high accuracy over large areas.

The difference in point densities from nadir to large scanning angles is related to the sinusoidal scanning pattern of the sensor systems used in the New Nationwide Elevation Model of Sweden. Using sinusoidal distribution, the distance between each point near nadir and on the edge of the swath are different (Baltsavias, 1999), with a consequent modification of the sampling density of the laser returns for different scanning angles.

Many of the observations that have been made in this investigation require further testing and sensitivity analysis considering the forest canopy structure and species composition, as well as the properties of laser pulses and scanning patterns. However, care must be taken when performing forest change detection using ALS-derived vegetation metrics related with the point

density of returns, such as the Understory and Vegetation ratios. The difference in magnitude of forest changes may be confused with apparent differences at coincident survey times due to off-nadir scanning angle effects.

The height percentiles were not greatly affected by the scan angle, which is in accordance with the findings of Holmgren et al, 2003). For the Southern site, the lower percentiles were more affected compared with the higher percentiles. For the Northern site, high and low height percentiles were affected to about the same degree. For both test sites, the middle height percentiles were least affected.

As mentioned earlier, the Vegetation ratio was affected most by the scan angle effect for both test sites. However, Cumulative Density was not affected as much. Thus, some of the errors for biomass or volume estimates obtained in earlier studies could be caused by the inclusion of the Vegetation ratio. Thus, another solution to reduce the scan angle effect might be to improve the selection of variables considering the scan angle effect.

References

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing*, 33, pp. 110-117.

Evans, J. S., Hudak, A. T., Faux, R. and Smith, A. M., 2009. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sensing*, 1, 776-794.

Holmgren, J., Nilsson, M. and Olsson, H., 2003. Simulating the effects of lidar scanning angle for estimation of mean tree height and canopy closure. *Canadian Journal of Remote Sensing*, 29, 623-63.

Hopkinson, C. and Chasmer, L., 2009. Testing LiDAR models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment*, 113, 275–288

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80, 88-99.

Nilsson, M., 1996. Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, 56, 1-7.

Parker, G. G. and Russ, E. M. 2004. The canopy surface and stand development assessing forest canopy structure and complexity with near-surface altimetry. *Forest Ecology and Management*, 189, 307-31.

Petersen, Y.M. and Burman Rost, H., 2011. Swedish Lidar Project - New Nationwide Elevation Model. *GIM International*, 2 (25), 21-23.

R Development Core Team, 2011. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing Access: http://www.R-project.org.

Toth, C. K., 2009. Strip Adjustment and Registration. In: J. Shan and C.K. Toth (Eds.), *Topographic Laser Ranging and Scanning – Principles and Processing*. CRC Press, Taylor & Francis Group: 235-268.

Vastaranta, M., I., Korpela, I., Uotila, A., Hovi, A. and Holopainen, M., 2011. Area-based snow damage classification of forest canopies using bi-temporal lidar data. In: D.D. Lichti and A.F. Habib (Eds.). *ISPRS Workshop Laser Scanning*, Volume XXXVIII-5/W12, Calgary, Canada: 5.

Table 1. Plot-level Wilcoxon signed-rank test for the Northern Site

Survey interval: 21-08-09 to 26-09-09	Scanning Angle classes (angles expressed in degrees and in absolute value)																				
	0-0	0-1	0-2	0-3	0-4	0-5	0-6	0-7	0-8	0-9	0-10	0-11	0-12	0-13	0-14	0-15	0-16	0-17	0-18	0-19	0-20
all returns (≥ 0 m)						**		***	***	***	***	***	***	***	***	***	***	***	***	***	***
ground returns (= 0 m)						**		*	*	*	**	*	**		.	***	***	***	***	***	***
all canopy returns (> 0 m)						.	.	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Understory returns (> 0 m to < 2 m)				.				*		**	***	***	***	***	***	***	***	***	***	***	***
canopy returns (≥ 2 m)						.	*	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Mean height (m)													*	*		.	**			**	.
Modal (m)									*				.	*			*				
Standard Deviation of height (m)																		*			
Median Absolute Deviation (m)										.							.			*	
Average Absolute Deviation (m)																		**			
Interquartile Range (m)																	*	.			
Skewness			.																		
Kurtosis							.			*		*			.		*			.	
10th percentile (m)														*	.		*	*	*	*	**
25th percentile (m)																	**			*	*
50th percentile (m)				.							.		**			*	*			*	
75th percentile (m)				.					.				*		.		.				
80th percentile (m)									.												
90th percentile (m)	*									.					.				**		
95th percentile (m)																**		.	*		
99th percentile (m)																**	.	***	***		
Maximum height (m)												.				**		***	***		
Cumulative Density of layer 1 (%)							.					*				*				***	*
Cumulative Density of layer 2 (%)															*						
Cumulative Density of layer 3 (%)													*			.	*	*			
Stratified Density of layer 1 (%)											*					.	*				
Stratified Density of layer 2 (%)			*				*						*			*	**	*	.	*	
Stratified Density of layer 3 (%)			.				*				.		**			*	***	*	.	***	.
Understory ratio (%)			.							**	***	***	***	***	***	***	***	***	***	***	***
Vegetation ratio (%)		**					.		**	.	**	***	***	***	***	***	***	***	***	***	***
Canopy Relief ratio index (%)			.							*											
Point Height Diversity index (%)		*									.	.	**	.	.	*	**	***		**	.
n. of plot	78	77	55	59	67	68	67	53	59	49	67	72	65	67	57	65	73	52	59	41	59

Significance levels: *** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05; . p-value < 0.1; blank space = not significant.

Table 2: Plot-level Wilcoxon signed-rank test for the Southern Site

Survey interval: 02-04-2010 to 01-07-2010	Scanning Angle classes (angles expressed in degrees and in absolute value)																				
	0-0	0-1	0-2	0-3	0-4	0-5	0-6	0-7	0-8	0-9	0-10	0-11	0-12	0-13	0-14	0-15	0-16	0-17	0-18	0-19	0-20
all returns (≥ 0 m)							*	*	***	***	**	***	***	***	***	***	***	***	***	***	***
ground returns (= 0 m)	*		*											*	**	**	***	***	***	***	***
all canopy returns (> 0 m)							*	***	***	***	***	***	***	***	***	***	***	***	***	***	***
Understory returns (> 0 m to < 2 m)		*							**	**	***	***	***	***	***	***	***	***	***	***	***
canopy returns (≥ 2 m)							*	***	***	***	**	***	***	***	***	***	***	***	***	***	***
Mean height (m)					.					*		.	*	**	*	*					**
Modal (m)															.						
Standard Deviation of height (m)					.							.			.	.					
Median Absolute Deviation (m)												*	*							.	
Average Absolute Deviation (m)	.								*			.		.	.			.		*	
Interquartile Range (m)									*			*						*		*	.
Skewness	*	*													**						**
Kurtosis									.	*		*			***		.		**		***
10th percentile (m)	.										.	.	*	**	*	*				*	*
25th percentile (m)												**	*	*	.			*		*	**
50th percentile (m)		*								.		*		*		*			.		.
75th percentile (m)											.		**	***							
80th percentile (m)							.						.	*							
90th percentile (m)								*						.							
95th percentile (m)			*																		
99th percentile (m)			.										*						*		
Maximum height (m)													**						*		
Cumulative Density of layer 1 (%)		.			.										.			*	.		
Cumulative Density of layer 2 (%)				.						.				*							
Cumulative Density of layer 3 (%)					*								*	*							
Stratified Density of layer 1 (%)				.				.					*	.		**			.	.	.
Stratified Density of layer 2 (%)		.								*				*							
Stratified Density of layer 3 (%)				.	.					**		*	*	**					*		
Understory ratio (%)		*							**	**	***	***	***	***	***	***	***	***	***	***	***
Vegetation ratio (%)						*		***		**		***	***	***	***	**	**	***	***	***	***
Canopy Relief ratio index (%)	.	.											.		**						*
Point Height Diversity index (%)					.			.				*	*	.	.	.			.	*	.
n. of plot	57	46	57	37	44	55	52	37	46	54	36	47	57	50	57	42	50	43	45	41	48

Significance levels: *** p-value < 0.001; ** p-value < 0.01; * p-value < 0.05; . p-value < 0.1; blank space = not significant.

A voxel-based simulator of lidar point clouds for improving forest attributes extraction methods

Brindusa-Cristina Budei [1], Benoît St-Onge [2] & Chhun-Huor Ung [3]

[1]Institut des sciences de l'environnement, Université du Québec à Montréal,
budei.brindusa_cristina@courrier.uqam.ca
[2]Département de géographie, Université du Québec à Montréal,
st-onge.benoit@uqam.ca
[3]Canadian Forest Service, Natural Resources Canada,
Chhun-Huor.Ung@RNCan-NRCan.gc.ca

Paper Number: SL2012-133

Abstract

Two major types of approaches for predicting forest attributes from remote sensing data can be schematically distinguished: plot or stand area-based approaches, and individual tree approaches. Both need calibration and verification data for their development and application. This requires expensive field sampling of a large number of plots to obtain estimates of forest attributes with sufficient accuracy and representativeness. What is more, these models must be recalibrated for every new airborne survey due to the variation in lidar point density.

To address these problems and to facilitate the improvement of current methods, we designed a simple simulator of lidar point clouds. This enables us to create controlled experimental conditions in which tree and stand attributes, as well as the virtual flight parameters of the lidar are modelled. The simulator consists of two modules: a voxel based forest module and a pulse generator module. The forest module generates a voxel representation of trees whose attributes are defined by species-specific allometric relationships (tree height, crown shape, crown length, crown width and plant area index). The number, location and attributes of voxelized trees can be read from existing field records or automatically generated in order to reproduce certain stand level attributes (number of trees per hectare, percentage of species and diameter class). The tree crown voxels possess a value representing the probability it may reflect a pulse. The pulse generator module simulates the trajectory and speed of the aircraft, the pulse repetition rate and the scan frequency. Pulses are propagated through the stand canopy and then returns are generated.

The development and application of the simulator is being carried out in three steps: (1) validation of the simulator with the real world data, (2) application of the simulator to create sets of point clouds through which the forest parameters and lidar parameters can be changed, and (3) evaluation and improvement of the stand and individual tree attributes extraction methods. This presentation focuses on the results of the first two steps.

For validating the simulator based on real plots and lidar data, we reproduced the trees previously measured and geo-located in 20 plots of 400 m2, and we then generated returns based on the recorded aircraft trajectory. GPS time was used to associate the location of the true first return of each pulse with the corresponding sensor position. Then, the equation of the laser ray was obtained and used to propagate the virtual pulse through the voxel canopy. Statistical comparison (quantile-quantile and vertical point count profile) between resulting simulated point clouds and those of the actual flight allowed us to validate the simulator.

As a first experiment, we show how tree density, when height distribution and other parameters are held constant, is reflected in the point clouds and various classical area-based metrics (percentiles, etc.). This provides insight on the performance of current attributes extraction methods and guidance for their improvement.

1. Introduction

Two major types of approaches for predicting forest attributes from remote sensing data can be schematically distinguished: plot or stand area-based approaches (Næsset 2002; Maltamo, Malinen *et al*. 2006), and individual tree approaches (Morsdorf, Meier *et al*. 2004; Popescu 2007; Yu, Hyyppä *et al*. 2011). In the former, equations are developed to predict forest attributes at stand level from variables characterising distribution of height of lidar points. For the latter, equations are formulated for predictions of characteristics of individual trees, after a previous segmentation of lidar points cloud. Both need calibration and verification data for their development and application. This requires expensive field sampling of a large number of plots to obtain estimates of forest attributes with sufficient accuracy and representativeness. What is more, these models must be recalibrated for every new airborne survey due to the variation in lidar point density, flight pattern, etc.

The simulation of lidar surveys over forest stands is an emerging approach for testing the adequacy of the scan parameters or the accuracy of the analysed methods. We designed a simple simulator of lidar point clouds in order to facilitate the improvement of current methods for analysing lidar data. This enables us to create controlled experimental conditions in which tree and stand attributes, as well as the virtual flight parameters of the lidar, are modeled. Provided the simulations are sufficiently realistic, the simulator could be used to generate series of points clouds for various individual (virtual) plots according to the lidar acquisition specifications of a real survey. Calibration of area-based regressions for extracting forest attributes for this survey could then be based on these artificial points clouds, instead of the now classical field plot based approach.

The main difference between existing models is the tree representation method (i.e., raster or vector based). The raster models are less expensive in memory and calculation time and are mainly used to obtain statistically similar point cloud as the airborne laser scanning (ALS) using geometric shapes of crown, in order to test different analytical methods. The simplest raster tree crown models use plain geometric volumes without laser beam crown penetration (Sun and Ranson 2000; Lovell, Jupp *et al.* 2005; Frazer, Magnussen *et al.* 2011). Models that simulate the laser beam penetration through the crown assign a density value to each voxel following a certain probability or clumping degree and provide an algorithm for laser propagation (Goodwin, Coops *et al*. 2007). The vector representation of trees uses a very detailed tree models and employs ray tracing methods in order to test the interaction between laser ray and the canopy (Disney, Kalogirou *et al*. 2010).
We choose a raster-based approach to simulate tree crown and a ray casting algorithm to calculate the penetration of laser beam through the crown. As we will generated thousands of simulations at plot level, the computations must be very fast, hence we have made compromises between realism and efficiency. We also introduce a method for the validation of point cloud simulations based on the real lidar ray trajectory in order to limit the uncertainty sources when we compared simulated and real point clouds.

2. Methods

2.1 Simulation principles

The simulator consists of two modules: one that creates 3D models of plots at the individual tree level, and another for airborne laser pulse generation. The simulator generates 3D scenes using 25 cm voxels at the plot scale (400 m^2 in the present study). The forest module generates a voxel representation of trees whose attributes are defined by species-specific allometric relationships (tree height, crown length and crown width related to diameter at breast height (DBH). The crown shape is generated using a general and versatile equation, suggested by Horn (1971), which produces different shapes that can mimic the envelope of hardwood or softwood crowns.

$$x^a + (by)^a = c^a \quad (1)$$

where x is crown radius, y is the relative height of the voxel to the crown base height, a is the shape parameter, b is the height to width ratio, and c is the absolute size.

The tree crown voxels are given a value representing reflectance ($0 \leq R \leq 1$, with transmittance = $1 - R$). In the current version of the simulator, foliage voxels are randomly attributed a R value between 0 and 0.2, with normal distribution, and tree trunk voxels a R value of 1 (thus achieving zero transmittance). The ground layer is populated with voxels with a R value of 1, and atmospheric (void) voxels have a zero value of R. A pulse having an initial intensity I (above canopy) is propagated through the voxels. Returns are generated if RI exceeds an experimentally determined threshold. I is attenuated by $1-R$ twice (incident and reflected paths). Pulses travel along straight lines and have a beam diameter of zero. RI is calculated for all intersected voxels over all the travel length in which the pulse is not completely attenuated or blocked by a tree trunk or the ground. The spatial distribution of R values within crowns does not vary between species and does not model clumping or specific distribution of leaf density. Moreover, multiple scattering was not modelled. R values in voxels occupied by two overlapping crowns correspond to the maximum between the R values of the respective overlapping voxels of the different crowns involved. Returned intensity was sampled every 30 cm along the path of the ray.

Virtual plots are generated by planting 3D model trees filled with voxels on flat ground for simulated runs or on a true ALS digital terrain model (DTM) in the case of validation (see following section). The number, location and attributes of voxelized trees can be read from existing field records to reproduce real situations or automatically generated in order to attain certain plot level attributes (e.g. number of trees per hectare, percentages of given species, diameter distribution, etc.). Tree size and shape are determined based on diameter at breast height (DBH) and species using allometric models. Diameter distributions can be made to follow a uniform or Weibull distribution of a given a mode.

The pulse generator module simulates the trajectory and speed of the aircraft, the pulse repetition rate, the (zigzag) scan frequency and the maximum scan angle. It can replicate a real aircraft path by reading the trajectory file of a completed survey (*.trj* file in the case of Optech lidars), or follow straight and parallel simulated flight lines of given altitude and *XY* position having a regular spacing. Plots can receive pulses from different flight lines.

Both simulation modules were developed in C++ and run on a standard PC machine. It should be remebered that our modelling approach is a compromise between speed and realism as thousands of simulations of plots will be produced at a 0.25 m voxel resolution with a high density of incident pulses. For this reason, the modelling approach does not consider all possible physical phenomena of the canopy.

2.2 Validation of the simulator

The goal of the validation is to verify that the synthesized point clouds are sufficiently realistic to be used to guide the improvement of information extraction methods. It is not aimed at proving that the physical processes involved are precisely modelled. For this reason, validation is performed at a general level. Validation consists in using the simulator to reproduce real point clouds resulting from an ALS carried out in 2010 over mixed stands of a test area. The simulated points clouds are produced based on the aircraft trajectory files and the exact location and size of field measured trees.

The test area is situated in the Matane reserve (67°11'W, 48°37'N), in the province of Quebec, Canada. The predominant species are balsam fir (*Abies balsamea*), paper birch (*Betula papyrifera*), yellow birch (*Betula alleghaniensis*), trembling aspen (*Populus tremuloides*), eastern hemlock (*Tsuga canadensis*), eastern whitecedar (*Thuja occidentalis*), white spruce (*Picea glauca*), red maple (*Acer rubrum*) and pin cherry (*Prunus pensylvanica*). 20 sample plots of 400 m^2 were chosen to sample situations showing a variety of species, ages, densities and forest structure.

Standard dendrometric measurements were made *in situ*, comprising species identification and DBH measurement for all trees in the plot (DBH above 4 cm). Finer measurements were performed for the five biggest trees in each plot (tree height, base height, crown diameter and location). Relative position from the center of the plot (compass and vertex) was measured for all trees above 8 cm of DBH in 5 plots. Plots location was measured by GPS (Trimble) and absolute location of trees was then calculated with an expected error of approximately 3 m. In addition, terrestrial lidar (Z+F 5010) for the 20 plots data were collected in May 2010, in leaf off conditions for hardwood trees. Four or five scans per plot were registered depending on tree density. The scans were aligned using Z+F software, georeferenced and filtered in order to perform a precise correction of the positions of trees measured in the field.

Airborne lidar data were collected in June 2010 using an Optech ALTM3100. The mean density of all returns is 15 points/m^2. Up to four returns were recorded for each pulse. The airborne trajectory is formed by a series of XYZ GPS coordinate measured at a one second interval. The raw lidar data was processed with TerraSolid software for flight line alignment, and point classification. Returns were stored within LAS files. Table 1 presents the main flight and sensors specifications.

Table 1: Specification of the lidar data acquisition

ALTM Sensor	ALTM3100
Pulse repetition frequency (kHz)	100
Flight altitude (meters about ground level)	530 - 1050
Scan Rate (Hz)	66
Scan Angle (deg)	15
Approx. points per m^2	15

To validate the simulator we first reconstructed the trees measured and geolocated in the field in 400 m^2 plots. Fine correction of GPS geolocation of the center of the plot was performed by comparing the real lidar canopy height model (CHM) with the one generated from simulated trees reproducing field data. Slight and systematic translations in X and Y were applied to all trees to improve the overall, plotwise planimetric coregistration of the real and simulated trees. Tree geolocations relative to the plot center were then individually corrected using the terrestrial lidar data as a reference. Modeled tree sizes and shape were determined using their true DBH

and species-specific allometric equations, and manual height measurements performed on the aerial and terrestrial lidar data. Trees below the threshold DBH of 8 cm and shrubs were not modelled. We used the Digital Terrain Model (DTM) obtained by interpolation of classified ground returns from the real lidar data as the bare earth elevation in the simulation.

We then generated returns based on the true aircraft trajectory recorded in flight. Using GPS time in the trajectory and LAS files as a key, we linked the XYZ origin of each pulse with the observed XYZ position of the corresponding first return. The ray equation was thus calculated and used to guide the propagation of simulated pulses through the voxel representations of field trees.

The simulated and corresponding real lidar were compared using three methods: 1) visual comparison of the real and simulated point clouds, 2) comparison between vertical point count profiles by 1 m strata of the real and simulated point clouds (using the R^2), and 3) a comparison of certain metrics of the vertical distribution of returns (e.g. difference between corresponding height quantiles of the real and simulated point clouds).

2.3 Generation of series of simulations for forest structure analysis

We define "simulation series" as a collection of realizations (realization = one model run for a given plot and a given set of parameter values) in which one or more forest attributed are varied in regular increments between a minimum and maximum values while other attributes are held constant. At each increment, different realizations are computed in which only the location of the trees within the plot space is changed. Lidar pulses are generated from linear flight lines and their pattern is determined by the simulated survey and sensor characteristics (see Table 2 and 3). Metrics are then extracted from each of the realization and regressed, for a given series, against the forest attribute of interest. As a first experiment, we show how tree density, when height distribution and other parameters are held constant, is reflected in the point clouds and various classical area-based metrics (percentiles, etc.). The density range between 100 and 1000 trees/ha in a simulated plot of 650 m^2. The model can plot trees over an existing georeferenced terrain model or in a flat Cartesian space. We choose the latter to analyse the density effect on lidar point distribution. For a first simulation, we choose a crown shape which fit a coniferous crown. Height distribution follow a Weibull curve with a specified modal height (we choose 15, 20 and 25 m modal heights). The series specifications are presented in Table 2 along and the simulated lidar parameters in table 3. As this is a first exploration, we have varied only one parameter at a time to fully isolate its effect. This creates conditions which may not always reflect natural forest structures. For example, density and height are normally inversely proportionate. In future studies, we will also generate series with inter-parameter covariation.

Table 2 : Simulated stand characteristics

Series	Variable of interest : Density (trees/ha)	Controlled variable: mean height (meters)
1	Min : 100, Max : 1000 Increment : 100	15
2		20
3		25

Table 3 : Specification of the simulated lidar data

Pulse repetition frequency (kHz)	33

Flight altitude (meters about ground level)	1000
Scan Rate (Hz)	66
Maximum scan angle relative to nadir (deg)	15
Approx. total returns per m^2	15
Aircraft speed	70 m/s
Scan pattern	zigzag

For each realization, the following metrics were extracted: 1st, 2nd and 3rd quartiles (*Q1*, *Q2*, *Q3*), 95th percentile (*P95*), and proportion of first returns respectively below 1 meters (*D1)*, without any minimum height threshold being applied. These are among the most commonly used metrics and are part of the set proposed by the FUSION software package of the USDA. Metrics extraction was carried out using special purpose software developed in C++.

3. Results

3.1 Validation of the simulator

A first appraisal of the simulation approach is based on a visual analysis of a plot of simulated trees reproducing the geolocation and size of trees of one field-measured plot (Figures 1 & 2). Hardwood and softwood can be distinguished by their crown shape.

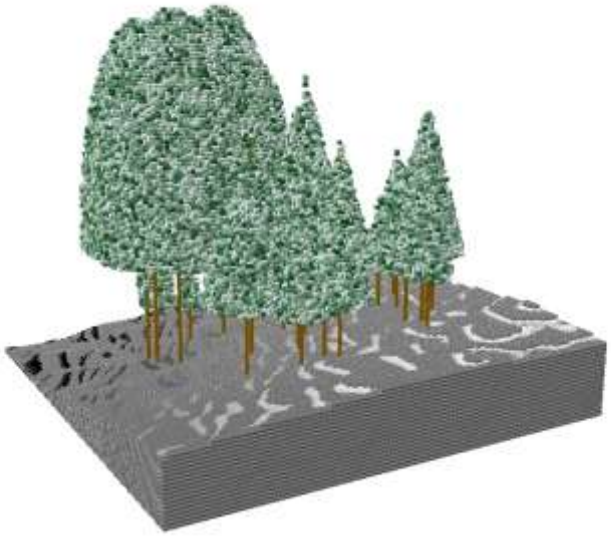


Figure 1: Exemple of a voxel-based simulated plot. Green intensities in tree crown represent random values of reflectance.

The simulated lidar returns are compared to the corresponding the real ones in Figure 2. Color coding of the return number helps in comparing the two point clouds. These can be said to be visually very similar. Crowns appear better defined in the simulated point cloud. Moreover, as we did not simulate understory vegetation nor trees with DBH smaller than 8 cm, more ground returns are present in the simulated point cloud than the real lidar data, specifically for first returns.

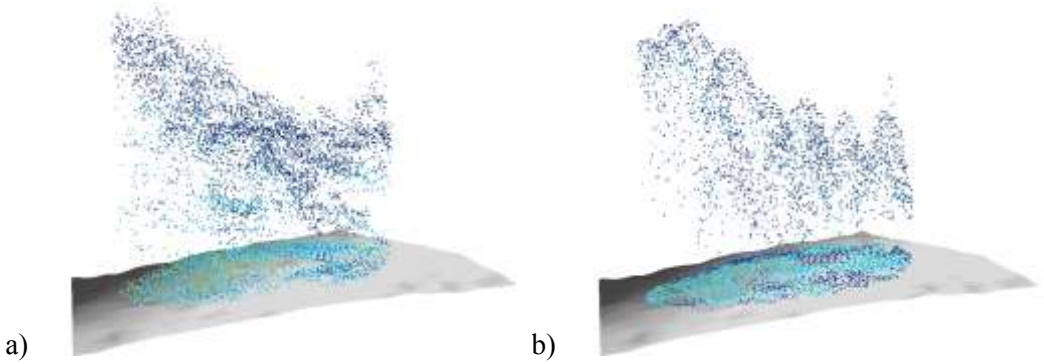


Figure 2 : a) Real lidar returns; b) Corresponding simulated lidar points. Point colour represents return number (first return –dark blue, second return – light blue, third return - green). The real lidar DTM appears in shades of grey.

The simulations were further verified by comparing the distribution of number of point by strata of 1 m for lidar points and simulated points (respectively for the total number of returns, first, second and third returns). Linear regressions between these curves give an overall R^2 of 0.97 for all the returns, of 0.54 for the first, of 0.98 for the second and 0.99 for the third returns (Figure 3). As we did not simulate the understory and trees with DBH less than 8 cm, the differences between curves at ground level are not considered fully indicative of model performance.

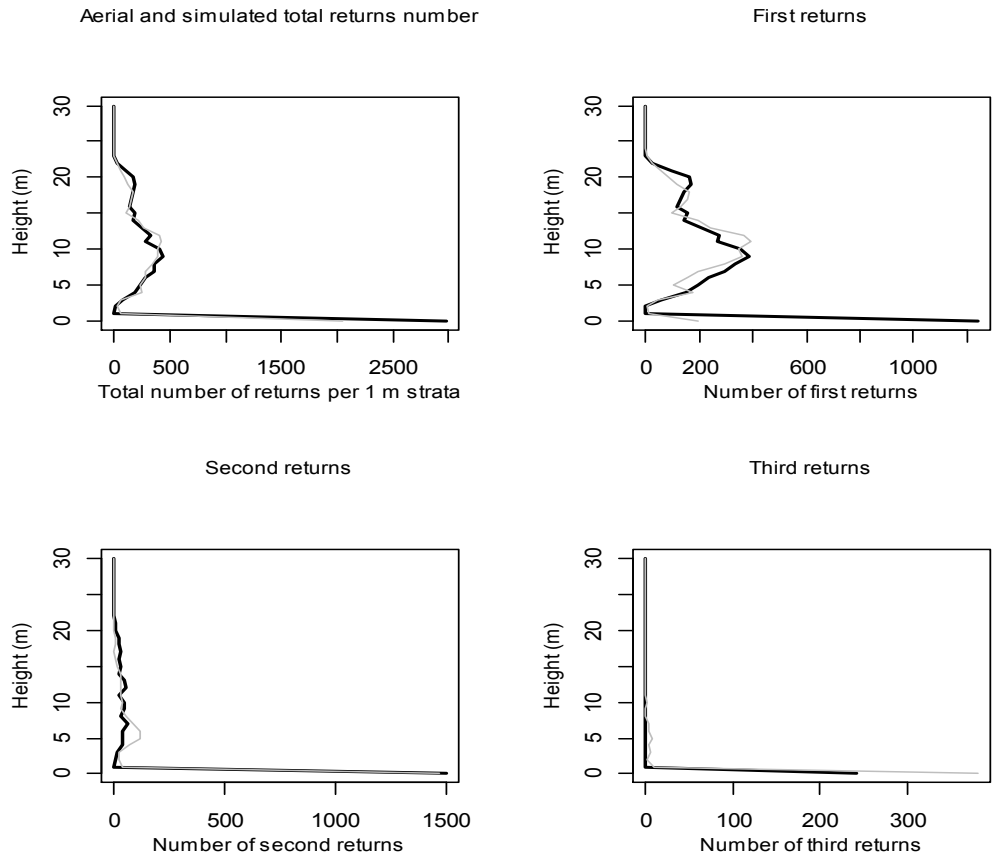


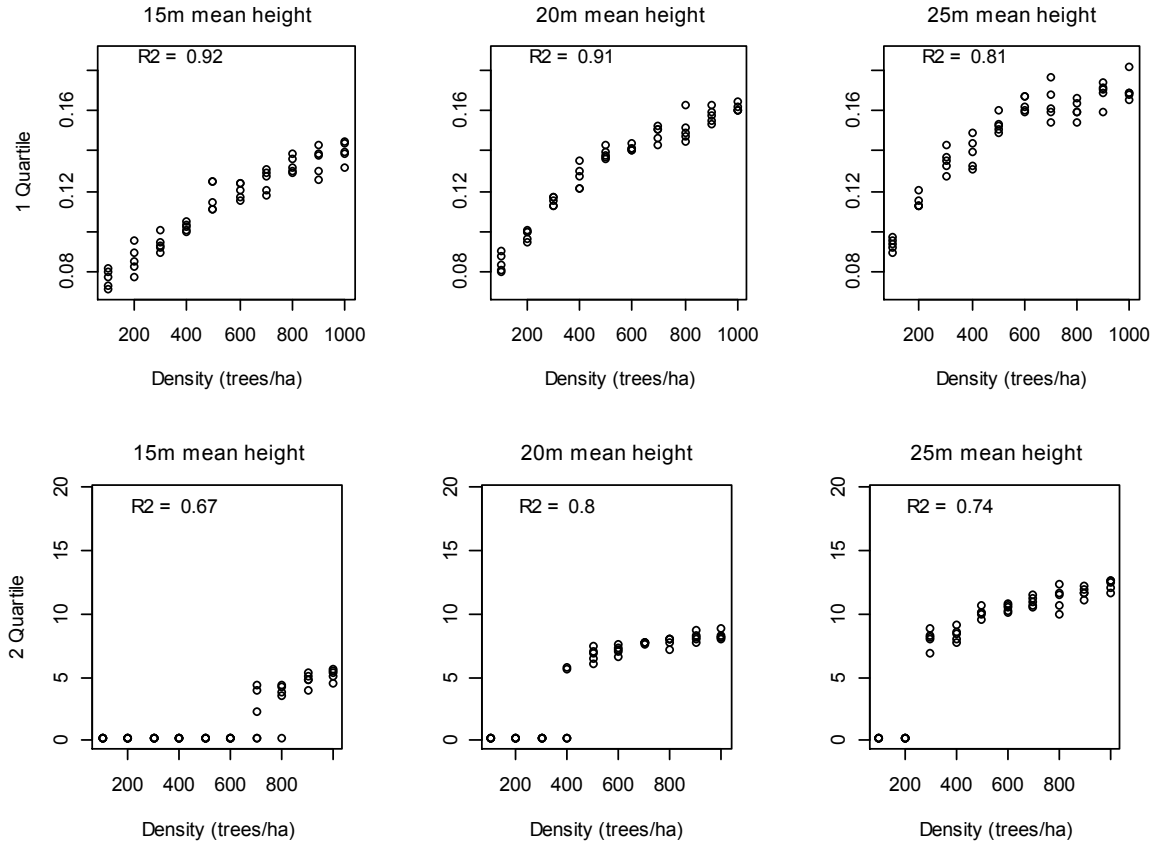
Figure 3: Comparison between profile of simulated (black line) and real lidar data (gray line).

The strong positive correlation between simulated and real lidar profiles curves is satisfactory as

well as the high similarity between the simulated and real CHMs may be considered sufficient for deriving initial conclusions from the simulations of series. At the same time, the model is dependent on the within crown distribution of reflectance values, the sampling density along the laser rays and thresholds in the algorithm of discretization of laser returns. The realism of crown shape representation is an other important factor. To improve it, species specific parameters must be considered in the future. Before more data series are generated, further improvements will be made to the simulator.

3.2 Series of simulations

A total of three series were generated in which either the density, the height or the crown shape was varied. We first present a subset of graphs showing the relationship between forest density and selected lidar point cloud metrics. We report the coefficient of determination (R^2) of the linear regression (Figure 4 and 5).



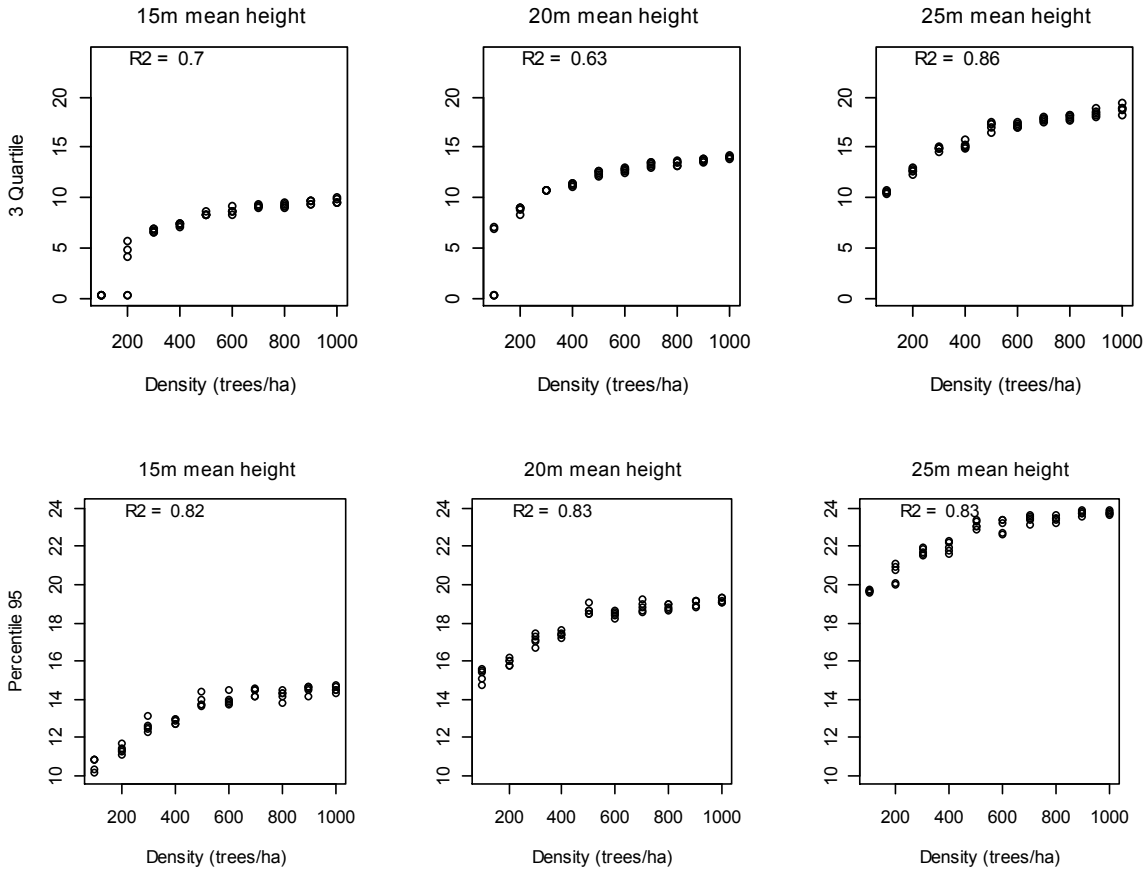


Figure 4: Variation of 1st, 2nd and 3rd quartiles (in m) with tree density for different modal tree heights.

The graphs of the quantiles (25, 50, 75 and 99, in Figure 4) reveal strong correlation with tree density, but show that these relationships are non linear. The overall shape of the relationship is not affected by average tree height. We see that the quantile values for different realizations vary to some degree showing the even for a given set of precisely defined input parameters (density, tree shape and height distribution), the precise location of trees influences the resulting quantile value and contributes to the overall point scatter.

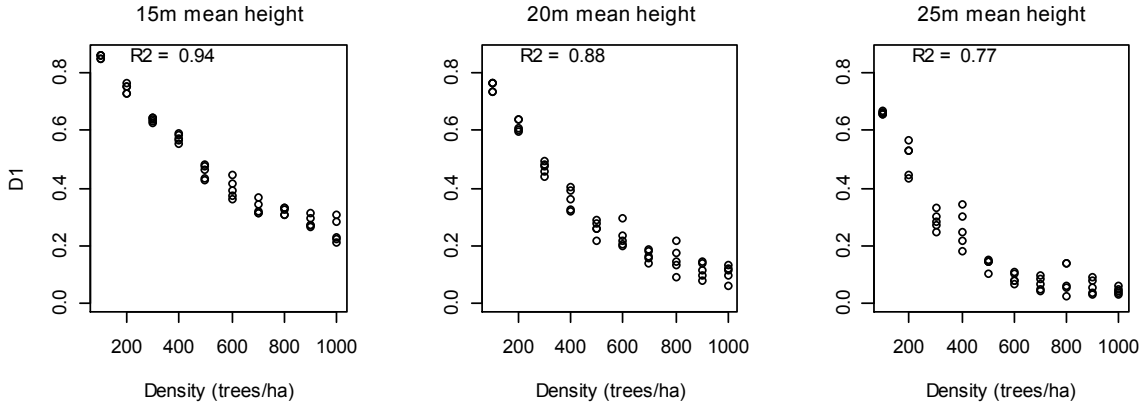


Figure 5 : Variation of proportion of first returns above 1 m with tree density for different mean tree heights.

The relationship between D1 and tree density (Figure 5) shows a stronger correlation for smaller trees. Its shape is also nonlinear.

4. Conclusion

Following up on the work of previous studies, we have developed, tested and used a new lidar point cloud simulator in which crown penetration and multiple returns are modelled. For the first time (to the best of our knowledge) such a simulator was validated using the exact laser rays trajectory of a real survey. Initial results are very encouraging as they highlight the fact, for example that, in controlled conditions, the relationship between the density of trees and certain quantiles are not linear, thus justifying the use of log-transformed variables. After final improvements and validation of the lidar simulator, we will use it to generate series reflecting numerous scenarios, such as pure and mixed stands, different lidar return densities, covariation of structural attributed, etc. Observations on the variation of laser derived variables (percentiles of the laser canopy height, cumulative proportions of lidar returns, etc.) with respect to different characteristics of the simulated tree population should improve the understanding, calibration and standardisation of area-based or individual tree approaches.

Acknowledgements

We acknowledge Dr. Chris Hopkinson and AGRG team for the airborne lidar survey and base processing of data and Dr. Richard Fournier and the Sherbrooke University team for the terrestrial lidar acquisition.

References

Disney, M. I., V. Kalogirou, et al., 2010. Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. *Remote Sensing of Environment* 114(7): 1546-1560.

Frazer, G. W., S. Magnussen, et al., 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment* 115(2): 636-649.

Goodwin, N. R., N. C. Coops, et al., 2007. Development of a simulation model to predict LiDAR interception in forested environments. *Remote Sensing of Environment* 111(4): 481-492.

Horn, H. S., 1971. *The adaptive geometry of trees*. Princeton, N.J., Princeton University Press

Lovell, J. L., D. L. B. Jupp, et al., 2005. Simulation study for finding optimal lidar acquisition parameters for forest height retrieval. *Forest Ecology and Management* 214(1-3): 398-412.

Maltamo, M., J. Malinen, et al., 2006. Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data. *Canadian Journal of Forest Research* 36(2): 426-436.

Morsdorf, F., E. Meier, et al., 2004. LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management. *Remote Sensing of Environment* 92(3): 353-362.

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* 80(1): 88-99.

Næsset, E., 2007. Airborne laser scanning as a method in operational forest inventory: Status of

accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research* 22(5): 433-442.

Popescu, S. C., 2007. Estimating biomass of individual pine trees using airborne lidar. *Biomass and Bioenergy* 31(9): 646-655.

Sun, G. R. and J. K. Ranson, 2000. Modeling lidar returns from forest canopies. *IEEE Transactions on Geoscience and Remote Sensing* 38(6): 2617-2626.

Treitz, P., K. Lim, et al., 2012. LiDAR Samplong Density for Forest Resource Intentories in Ontario, Canada. *Remote Sensing* 4: 830-848.

Yu, X., J. Hyyppä, et al., 2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(1): 28-37.

Predicting Live and Dead Tree Basal Area in Bark Beetle-Affected Forests from Discrete-Return LiDAR

Andrew T. Hudak[1], Ben Bright[2], Jose Negron[3], Robert McGaughey[4], Hans-Erik Andersen[5], Jeffrey A. Hicke[6]

[1]U.S. Forest Service, Rocky Mountain Research Station, ahudak@fs.fed.us
[2]University of Idaho, Department of Geography, bright@uidaho.edu
[3]U.S. Forest Service, Rocky Mountain Research Station, jnegron@fs.fed.us
[4]U.S. Forest Service, Pacific Northwest Research Station, bmcgaughey@fs.fed.us
[5]U.S. Forest Service, Pacific Northwest Research Station, handersen@fs.fed.us
[6]University of Idaho, Department of Geography, jhicke@uidaho.edu

Paper Number: ###SL2012-139

1. Introduction

Recent bark beetle outbreaks in western North America have been widespread and severe. High tree mortality due to bark beetles affects the fundamental ecosystem processes of primary production and decomposition that largely determine carbon balance (Kurz *et al*. 2008, Pfeifer *et al*. 2011, Hicke *et al*. 2012). Forest managers need accurate data on beetle-induced tree mortality to make better decisions on how best to remediate beetle-affected forests and restore healthy ecosystem services (Negron *et al*. 2008). Discrete-return LiDAR measures canopy height and density, and LiDAR intensity provides some indication of the spectral reflectance and condition of canopy elements (foliage and branches) (Kim *et al*. 2009). LiDAR has been successfully applied to estimate biomass and carbon stocks in healthy forest (Hudak *et al*. 2012) and beetle-affected forest (Bright *et al*. 2012). A challenge in beetle-affected forests is that most airborne LiDAR has a single near infrared wavelength; i.e., LiDAR lacks the multispectral information useful for distinguishing between green, red, and grey trees. However, LiDAR intensity values may help distinguish between live green and dead red or grey trees. Moreover, mountain pine beetles (the most widespread bark beetle currently) and spruce beetles preferentially attack larger trees, so beetles impart a canopy structural signature that may be exploited (Coops *et al*. 2009).

Our objective is to predict Live and Dead Basal Area (BA) in beetle-affected areas in five states in the USA using canopy height, density, intensity, and topographic metrics derived from discrete-return airborne LiDAR data, tree measurements collected in field plots and summarized into plot-level estimates of Live BA and Dead BA using the Forest Vegetation Simulator (FVS), and the nonparametric Random Forests (RF) machine learning algorithm (Breiman 2001). Predicting both Live and Dead BA in bark beetle-affected forest, where live and dead trees are typically thoroughly mixed, has not been attempted before and should provide insight into the sensitivity of LiDAR to bark beetle effects on coniferous forest canopies.

2. Methods

2.1 Study areas

Areas of bark beetle-infested coniferous forests included in this study are from: 1) the Kenai Peninsula in Alaska (AK), 2) the Pinaleño Mountains of southeastern Arizona (AZ), 3) north-central Colorado (CO), 4) central Idaho (ID), and 5) central Oregon (OR). The commonality among these study areas is that both field plot and LiDAR data were collected at

the time of an active beetle outbreak in mature coniferous forest with each study area having trees of variable age, height, and density. The range of bark beetle-induced mortality was also highly variable across the field plots, ranging from 0 to 100% mortality.

2.2 Field plots

Field plot locations in AZ, CO, and ID were selected using stratified random (CO and ID) or systematic (AZ) sampling designs within the bounds of the LiDAR acquisitions. All plots were of fixed radius with radii of 12.6 m in AZ (0.05 ha), 8.0 m in CO (0.02 ha), and 11.3 m in ID (0.025 ha). Field plots in AK and OR were Forest Inventory and Analysis (FIA) plots at predetermined, systematic locations. In OR, these plots were augmented with USFS Region 6 (R6) Continuous Vegetation Survey (CVS) plots also situated along a systematic grid. Each FIA plot in AK and OR consists of 5 subplots of 7.3 m radius (0.11 ha total plot area) while the R6 plots in OR consist of 4 subplots of 15.5 m radius (0.38 ha total plot area). The subplot level data were combined and summarized at the plot level to preclude any issues of spatial autocorrelation, pseudo-replication, or inflated plot numbers. Live and Dead BA were summarized at the plot level from individual tree diameter at breast height (dbh) measurements and tree condition calls using the Forest Vegetation Simulator (FVS) while taking into account the different plot areas and then converted from English units of ft^2/ac to metric units of m^2/ha.

2.3 Data processing

Metrics of the height, intensity, and density of LiDAR points within the fixed-radius field plots were calculated using the CloudMetrics tool of FUSION LiDAR visualization and analysis software (McGaughey 2012). In AK and OR, the LiDAR points within the individual subplots were extracted from the point cloud and concatenated into a single file prior to calculating plot-level metrics. In all study areas, ground elevations from the LiDAR-derived digital terrain model (DTM) were subtracted from the point elevations before calculating canopy metrics. A height threshold of breast height (1.37 m) was used to separate ground and understory returns from overstory returns. Height and intensity metrics were statistics based on the distributions of height and intensity values associated with each return: percentiles (minimum, median, maximum, quartiles and interquartile range (IQ)), mode, mean, standard deviation, variance, skewness, kurtosis, coefficient of variation (CV), average absolute deviation (AAD), and canopy relief ratio (CRR; Pike and Wilson 1971). Density metrics were calculated as the percentage of both 1st and all returns >1.37 m (breast height, BH), mean, and mode; and within seven vertical height strata 0-6: <0.15m, 0.15-1.37 m, 1.37-5 m, 5-10 m, 10-20 m, 20-30 m, and >30m. Twenty topographic metrics, including elevation, slope, transformed aspect, profile and planiform curvatures, slope*cos(asp), slope*sin(asp), and heatload index were also calculated from digital terrain models (DTM) after aggregating them to 30-m resolution, using an Imagine add-on tool developed at the USFS Remote Sensing Applications Center (Ruefenacht In Prep).

2.4 Random Forests model development

Predictor variables were selected from the suite of candidate metrics generated from either FUSION (canopy height, intensity, density metrics) or the Imagine tool (topographic metrics). RF calculates variable importance values based on the number of nodes it forms within the random forest of regression trees. Those predictors exhibiting the highest Model Improvement Ratio (MIR) calculated from standardized importance values were selected (Evans and Cushman 2009, Evans *et al*. 2010). RF randomly selects variable subsets through a bootstrapping procedure that makes it robust to highly correlated data and resistant to overfitting (Breiman 2001). However, in the interest of parsimony, we also eliminated predictors if they had a Pearson correlation $r > 0.9$ with another predictor having a higher MIR value. RF randomly withholds one third of the data as an out-of-bag sample at every bootstrap iteration, so

it is not necessary to partition the data into model and validation datasets (Breiman 2001).

3. Results

In the AK, AZ, and OR field plots, dead trees comprised less than half of the BA measured and were a low proportion of Total BA in AK and AZ, whereas in CO and ID, more than half of the plot-level BA was from dead trees (Table 1).

Table 1: Summary Statistics of Plot-Level Live and Dead Basal Area (m^2/ha)

Study area	Plots	Response	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
Alaska	194	Live BA	0.0	4.1	8.6	12.6	18.7	62.3
		Dead BA	0.0	0.0	1.5	3.9	5.7	24.5
Arizona	101	Live BA	0.0	9.3	32.8	35.5	57.0	126.3
		Dead BA	0.0	4.9	10.0	16.9	23.3	103.8
Colorado	113	Live BA	0.0	9.7	17.2	23.1	30.7	110.6
		Dead BA	0.0	36.8	50.8	54.9	70.6	126.7
Idaho	27	Live BA	5.0	8.7	14.3	15.3	21.2	29.5
		Dead BA	3.0	11.6	17.6	18.6	23.6	43.7
Oregon	190	Live BA	0.0	10.4	16.6	20.3	28.1	99.9
		Dead BA	0.0	0.0	1.4	3.3	4.3	27.2

Metrics of all types (height, intensity, density, and topographic) were selected based on MIR for the RF models (Table 2). In general, height and density metrics were most important in predicting Live and Total BA, whereas intensity metrics were most important in predicting Dead, %Live, and %Dead BA. An exception was OR, where intensity metrics were not important, perhaps because the dead trees comprised the lowest proportion of Total BA of any study area. Topographic metrics were generally the least important metrics in all models. Because intensity values could not be normalized across all study areas, we also tested the effect of omitting intensity metrics from the models; the result (not shown) was more density metrics selected in place of intensity metrics, but generally a lower number of selected predictors overall, and lower variance explained in all study areas except OR. Intensity metrics were more important in explaining Live BA in CO and ID, where tree mortality was higher, than in AK, AZ and OR. In AK, AZ, CO, and ID, RF Models predicting %Live and %Dead BA were almost identical in terms of important variables (Table 2) and variance explained (Table 3), which is not surprising because these responses are mutually dependent. In OR, where sampled tree mortality was lowest, RF models predicting %Live and %Dead BA were not identical.

In all study areas, the RF models explained more variance in Live BA than Dead BA (Table 3). In AK, AZ, CO, and ID, intensity metrics were very effective predictors of %Live and %Dead BA; omitting intensity metrics (not shown) reduced variance explained, especially for predicting dead BA. Models of %Live and %Dead BA explained more variance (51.2- 74.0% variance explained) than models predicting absolute Live, Dead, and Total BA (17.1- 67.9% variance explained) in these areas. In OR, where tree mortality was lowest and intensity metrics were comparatively unimportant, %Live and %Dead BA models explained less variance (13.6-24.5% variance explained) than RF models predicting absolute Live and Dead BA (27.3-59.1% variance explained). In AK, CO, ID, and OR, RF models predicting Total BA explained more variance than RF models predicting Live and Dead BA. The variance explained by RF for predicting Total BA in AZ was intermediate to that of predicting Live and Dead BA. In all study areas, omitting intensity metrics (not shown) did not greatly affect variance explained by RF

models predicting Total BA.

Table 2: LiDAR Metrics Selected for RF Models Based on MIR and Pearson r <0.9

Study area	Response	Height	Intensity	Density	Topographic
Alaska	Live BA	Mode, 10th	CV	%1st>Mean, Strat0	Elev
	Dead BA	Max, CRR	Mean, StDev, CV	Strat4	Elev
	Total BA	Max, Mode, 30th	Max	%1st>Mean, Strat0, Strat4	Elev
	%Live BA	Mean, StDev, CV, Skew, CRR	Mean, StDev, CV, Skew, Kurt	%1st>Mean, Strat4	Elev
	%Dead BA	Mean, StDev, CV, Skew, CRR	Mean, StDev, CV, Skew, Kurt	%1st>Mean, Strat4	Elev
Arizona	Live BA	Mean, Mode, CV, 10th, CRR	StDev, Skew	%All>BH, Strat0, Strat2	Elev
	Dead BA	Mean, Mode, CV, IQ, Skew	Max, Mean, Skew, Kurt	%1st>BH, %All>Mode, Strat1	Elev, SlpCosAsp
	Total BA	Mean, Mode, CV, 10th, CRR	StDev, Kurt	%1st>BH, %All>Mode, Strat0, Strat2, Strat3, Strat4	Elev
	%Live BA	Mode, CV, 5th, CRR	Mean, Skew, Kurt, AAD	%1st>BH	Elev
	%Dead BA	Mode, CV, 5th, CRR	Mean, Skew, Kurt, AAD	%1st>BH	Elev
Colorado	Live BA	Max, Mean, Mode, Skew, 5th	Mean, Kurt	%1st>BH, Strat2, Strat4	Heatload, Slope
	Dead BA	Mean, 99th	Skew	%1st>BH, %1st>Mode, Strat4	Elev
	Total BA	Max, Mean, 25th	Max, Mean	%All>BH, %1st>Mode, Strat4, Strat5	Elev
	%Live BA	Min, Kurt	Mean, CV, Skew, Kurt	-	Slope
	%Dead BA	Min, Kurt	Mean, CV, Skew, Kurt	-	Slope
Idaho	Live BA	Max, CV, 1st	CV, IQ, Kurt	%All>BH, %All>Mode, Strat1, Strat4, Strat5	Elev
	Dead BA	Mode, Skew, 30th,	Max, Mean, Skew, AAD	%1st>Mode, Strat1	Elev, Slope, SlpCosAsp
	Total BA	Min, Max, Skew, AAD, 40th, 75th	-	%All>Mean, Strat5	-
	%Live BA	1st	CV, IQ, Kurt	Strat1	Elev
	%Dead BA	1st	CV, IQ, Kurt	Strat1	Elev
Oregon	Live BA	Mean, 20th, 95th, CRR	Mean, Kurt, AAD	%All>BH, Strat4	Elev
	Dead BA	Mean, Kurt, 1st, 99th	Mean, CV	%All>BH, Strat6	Elev
	Total BA	Mean, CV, Kurt, AAD	Mode, CV, Kurt	%All>BH, Strat4	Elev
	%Live BA	Mean, Skew, Kurt, 95th, CRR	Mean, Mode, CV, Kurt	%All>BH, %All>Mean, Strat2, Strat3	-
	%Dead BA	Mode, CV, IQ, 1st, 99th, CRR	StDev, Skew, Kurt	%All>BH, %All>Mean, Strat2, Strat3	-

Table 3: % Variance Explained by RF to Predict Plot-Level Basal Area Responses

Study area	Response	Number of predictors	% Variance explained
Alaska	Live BA	6	67.2
	Dead BA	7	46.1
	Total BA	8	67.9
	%Live BA	13	59.7
	%Dead BA	13	59.1
Arizona	Live BA	11	50.2
	Dead BA	14	17.1
	Total BA	15	34.9
	%Live BA	10	74.0
	%Dead BA	10	73.4
Colorado	Live BA	12	44.3
	Dead BA	7	39.4
	Total BA	10	57.5
	%Live BA	7	51.6
	%Dead BA	7	51.2
Idaho	Live BA	12	58.0
	Dead BA	12	41.4
	Total BA	8	58.9
	%Live BA	6	67.7
	%Dead BA	6	67.7
Oregon	Live BA	10	59.1
	Dead BA	9	27.3
	Total BA	10	66.0
	%Live BA	13	24.5
	%Dead BA	13	13.6

4. Discussion

We attribute the lower variance explained for Dead BA compared to Live BA to greater penetration of the canopy by laser pulses after needles drop from dead trees, and thus a lesser amount of information than would be reflected from a green or red canopy of the same height. Tree condition calls (live or dead) at the time of field plot characterization did not coincide with the time of LiDAR acquisition; thus, we do not know whether dying or dead trees were green, red, or grey when the LiDAR was collected. Temporal mismatches between field and LiDAR data collections are likely greater in AK and OR, because only some FIA or CVS plots are characterized in a given year.

Based on this preliminary analysis, we conclude that it may be difficult to generalize a consistent subset of LiDAR metrics that work across these different affected forest ecosystems to predict bark beetle effects on BA. It is more likely that we may be able to generalize analytical approaches, if not specific metrics. In future work, we will include disturbance vectors derived from Landsat image time series in the analysis, to more accurately predict the BA response variables. Integrating temporal disturbance vectors from Landsat with airborne LiDAR metrics that capture fine spatial detail should prove a more powerful approach for predicting canopy structure effects of bark beetle-induced tree mortality.

5. Conclusion

This analysis tested the feasibility of predicting both Live and Dead BA in bark beetle-affected coniferous forests from discrete-return LiDAR. We found that LiDAR-derived height, intensity, density, and topographic metrics explained more variance in Live BA than Dead BA, and usually the most variance in Total BA, which we conclude is due to the greater sensitivity of LiDAR to needled trees than to grey trees.

Acknowledgements

Funding was provided by the Special Technology Development Program of the U.S. Forest Service Forest Health Protection division. Funding to the University of Idaho was through Joint Venture Agreement 11-JV-11221633-184.

References

Breiman, L., 2001. Random forests. *Machine Learning*, 45, 5-32.

Bright, B.C., Hicke, J.A. and Hudak, A.T. 2012. Estimating aboveground carbon stocks of a forest affected by mountain pine beetle in Idaho using lidar and multispectral imagery. *Remote Sensing of Environment*, 124: 270-281.

Coops, N.C., Varhola, A., Bater, C.W., Teti, P. Boon, S., Goodwin, N., and Weiler, M., 2009. Assessing differences in tree and stand structure following beetle infestation using lidar data. *Canadian Journal of Remote Sensing*, 35: 497-508.

Evans, J.S. and Cushman, S.A., 2009. Gradient modeling of conifer species using Random Forests. *Landscape Ecology*, 5, 673–683.

Evans, J.S., Murphy, M.A., Holden, Z.A. and Cushman, S.A., 2010. Modeling species distribution and change using Random Forests. In: C.A. Drew, F. Huettmann, Y. Wiersma (Eds.). Chapter 8 in *Predictive Modeling in Landscape Ecology*. Springer.

Hicke, J.A., Allen, C.D., Desai, A.R., Dietze, M.C., Hall, R.J., Hogg, E.H., Kashian, D.M., Moore, D., Raffa, K.F., Sturrock, R.N. and Vogelmann, J., 2012. Effects of biotic disturbances on forest carbon cycling in the United States and Canada. *Global Change Biology*, 18, 7-34.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J., Martinuzzi, S. and Falkowski, M.J., 2012. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Remote Sensing of Environment*, 123: 25-40.

Kim, Y., Yang, Z., Cohen, W.B., Lauver, C.L., and Vankat, J.L., 2009. Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. *Remote Sensing of Environment*, 113: 2499-2510.

Kurz, W.A., Dymond, C.C., Stinson, G., Rampley, G.J., Neilson, E.T., Carroll, A.L., Ebata, T. and Safranyik, L. 2008. Mountain pine beetle and forest carbon feedback to climate change. *Nature*, 452, 987-990.

McGaughey, R.J., 2012. FUSION/LDV: Software for LIDAR Data Analysis Visualization. USDA Forest Service, Pacific Northwest Research Station. 170 p.

Negron, J.F., Allen, K., Cook, B. and Withrow, J.R. Jr., 2008. Susceptibility of ponderosa pine, Pinus ponderosa (Dougl. Ex Laws.), to mountain pine beetle, Dendroctonus ponderosae Hopkins, attack in uneven-aged stands in the Black Hills of South Dakota and Wyoming USA. *Forest Ecology and Management*, 254: 327-334.

Pfeifer, E.M., Hicke, J.A., and Meddens, A.J.H., 2011. Observations and modeling of aboveground tree carbon stocks and fluxes following a bark beetle outbreak in the western United States. *Global Change Biology*, 17, 339-350.

Pike, R.J. and Wilson, S.E., 1971. Elevation relief ratio, hypsometric integral, and geomorphic area altitude analysis. *Bulletin of the Geological Society of America*, 82, 1079–1084.

Ruefenacht, B. (In Prep) Digital elevation model derivatives. USDA Forest Service Remote

Sensing Applications Center, Salt Lake City, Utah.

Monitoring aboveground biomass in the Yukon River Basin using multi-scale lidar

Birgit Peterson1, Kurtis Nelson2, and Bruce Wylie3

1 ASRC Research and Technology Solutions, contractor to the U.S. Geological Survey (USGS), Earth Resources Observation and Science (EROS) Center, Sioux Falls, SD bpeterson@usgs.gov
2 USGS, EROS, Sioux Falls, SD, knelson@usgs.gov
3USGS, EROS, Sioux Falls, SD, wylie@usgs.gov

Paper Number: SL2012-147

Abstract

Current climate predictions indicate that northern high latitude ecosystems will experience significant change over the next century and beyond. This will likely impact forest structure, including aboveground biomass (AGB), and carbon fluxes in boreal forest ecosystems. Also, fire frequency and severity are both expected to increase, leading to further impacts on forest disturbance and carbon fluxes. Characterizing and monitoring these changes in forest structure and carbon exchange is critical to understanding their impacts on ecosystem performance. In Alaska, with its expansive and inaccessible forests, remote sensing is integral to an active forest monitoring system. This study leverages field observations, airborne lidar data, and data from the spaceborne Geoscience Laser Altimeter System (GLAS) to generate tree and shrub AGB estimates for the Yukon River Basin (YRB) of Alaska. Using these observed estimates and metrics calculated from the airborne lidar data, regression models were developed to predict tree and shrub AGB for the airborne lidar study areas. The models were then applied to the full set of airborne lidar data to generate AGB maps for the two sites. The airborne lidar-derived AGB maps were then used to bridge the biomass predictions from the local field observations to a large-scale assessment for the YRB relying primarily on GLAS and Landsat data. To determine the utility of the GLAS-based ABG estimates for assessing biomass change, GLAS and Landsat data were also assimilated within the YRB to determine biomass loss in areas burned during the severe 2004 and 2005 fire seasons.

1 Introduction

Northern latitudes are experiencing changing climates at greater magnitudes than other regions of the globe (Soja *et al*. 2007). New climate patterns will likely affect the structure, distribution, and composition of boreal vegetation characteristics in northern latitudes (Soja *et al*. 2007, Barrett *et al*. 2011, Wolken *et al*. 2011). Furthermore, climate changes also influence the disturbance regime of boreal regions. Wildland fire is a significant cause of vegetation disturbance in many boreal regions. Recent trends indicate that wildland fires are occurring with increasing frequency and with greater severity in boreal Alaska (Kasischke *et al*. 2010). How this altered fire regime will interact with future climate change patterns and other sources of landscape disturbance such as insect infestation and storm damage is not yet well understood.

To understand the potential influence of a changing fire regime on future ecosystem-climate feedback cycles it is necessary to quantify biomass loss to fire (Kasischke and Hoy 2012). Remote sensing has been invaluable for quantifying aboveground biomass (AGB) over large, and particularly remote, areas. Lidar is especially suited for deriving canopy structure and AGB information (Lefsky *et al*. 2002). However, small-footprint, airborne lidar data collections, that can provide detailed information regarding canopy structure, are relatively sparse in boreal Alaska. In contrast, available spaceborne data are large-footprint and lack spatial continuity, but

provide a significant number of samples from which to derive vegetation structure metrics. Data collected by the spaceborne Geoscience Laser Altimeter System (GLAS) have been used for canopy height and AGB estimation (e.g., Lefsky *et al*. 2005, Boudreau *et al*. 2008, Selkowitz *et al*. 2012). Recent work using GLAS data from interior Alaska has demonstrated that GLAS parameters are useful indicators of burn severity (Goetz *et al*. 2010).

In this study we used both airborne and spaceborne lidar data to map AGB before and after severe fire seasons to quantify the effect of fire on AGB. The goals were to: 1) derive pre- and postfire maps of AGB using a combination of airborne lidar and GLAS data; 2) estimate fire-related AGB losses in the Yukon River Basin of Alaska for the years 2004 and 2005; and 3) relate these losses to fire severity. The 2004 and 2005 fire seasons were highly significant for interior Alaska in recent history because of the large area burned. Should the fire regime continue to change as predicted, such fire seasons may become more typical in the region.

2 Study Area

This study focuses on the section of the Yukon River Basin (YRB) within interior Alaska. This portion of the YRB encompasses approximately 52,500 km^2 (Figure 1). According to the National Land Cover Database (NLCD) almost 30% of the land area is covered by evergreen forest, approximately 7% by deciduous forest, and nearly 6% by mixed evergreen/deciduous forest. Almost 35% of the land cover consists of shrublands. Wetlands comprise almost 10% of the land cover, primarily in the delta region. The terrain varies from the flat plains of the Yukon River to steep terrain of the Brooks Range to the north and the Alaska Range to the south. According to the Monitoring Trends in Burn Severity (MTBS; Eidenshink *et al*. 2007) project, between 2004 and 2005 154 fires burned in the YRB, affecting over 4.5 million ha.

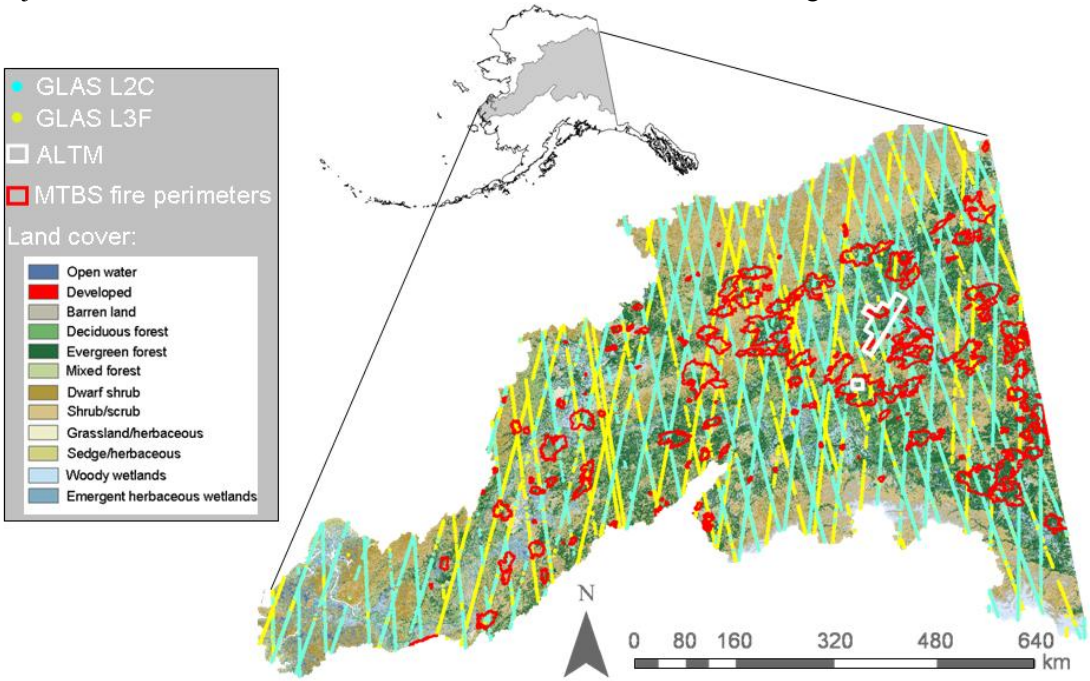


Figure 1 - YRB land cover with locations of data sets and 2004 and 2005 fire perimeters overlayed.

3 Data

3.1 Field sampling

Field data were collected in two study areas within the YRB, one in the Yukon Flats Ecoregion (YFE) and the other along Nome Creek (NC) in the White Mountains. These two areas

represent a range of vegetation and topographic conditions found in the YRB. The field data were collected in the summers of 2009 and 2010 at 70 plots. At each plot, breast-height stem diameters and/or basal diameters were measured which were used to derive AGB from species-specific allometric equations. The field data collection and the tree and shrub biomass AGB estimation are described further in Ji *et al*.(2012).

3.2 Airborne lidar

The airborne lidar data were collected in mid July and early September of 2009 in the YFE and in October 2012 for NC with an airborne Optec Airborne Laser Terrain Mapper (ALTM) Gemini and were delivered as bare-earth and first-return digital surface models. The point cloud data were delivered with returns labeled as either surface or canopy returns.

3.3 GLAS

GLAS GLA01 (waveform) and GLA14 (land/canopy elevation, footprint locations, fitted Gaussians) products were downloaded for two GLAS acquisition campaigns, L2c and L3f (both release 33). These acquisition times corresponded to leaf-on conditions within the study area. The L2c campaign collected prefire data, between May 18 and June 21, 2004. The L3f campaign collected data between May 24 and June 26, 2006, after the two significant fire seasons. A total of 73,359 GLAS footprints from L2c were located in the YRB and 103,031 from L3f.

3.4 Landsat

A mosaic of prefire Landsat data was developed from the Global Land Survey 2000 dataset, incorporating 66 Landsat 7 scenes from 1999-2002 (Bouchard *et al*. 2009). Postfire Landsat data were obtained from the Web-Enabled Landsat Data (WELD) project. WELD provides gridded seasonal and annual composites of Landsat 7 data over the conterminous US and Alaska for 2005 through 2011 (Roy *et al*. 2010). For the postfire image mosaic, a composite was created by selecting each pixel with the maximum NDVI from the 2006 through 2010 WELD annual images.

3.5 MTBS

Wildland fire perimeter and severity data were obtained from MTBS for all fires from 2004 and 2005 within the YRB. MTBS maps wildfire burn severity and perimeters across the U.S. using Landsat imagery from 1984 to present, and are widely used. Vector shapefiles of the perimeters were used to constrain the individual fire analyses and the classified burn severity images were used to compare with AGB change.

4 Analysis

4.1 ALTM-derived AGB

A set of metrics were derived from the ALTM point cloud data to use as dependent variables in a multiple linear regression analysis to predict AGB, using the field estimated AGB as the independent variable. The ALTM point cloud data were divided into 10 x 10 m cells. For each cell, the ground elevation was calculated as the average of all returns classified at ground returns. The height of each canopy return within a cell was calculated relative to the ground elevation of the given cell. The maximum, mean and covariance of the canopy heights derived from the lidar point cloud as well as deciles of cumulative canopy returns by height were also calculated. Additionally, density values for ten layers within the canopy portion of the point cloud were derived. The layers were defined by dividing the vertical distance between the elevation of the highest return and a 2 m elevation threshold. The densities for the individual layers were then derived by dividing the number of returns above the layer elevation to the total number of returns. Finally, within each grid cell the ratio of canopy returns above six height

thresholds (at 1 m, 2 m… 6 m above ground) to the total number of canopy returns was calculated.

Using the ALTM metrics and the field-estimated AGB values a regression model was developed to predict AGB from the ALTM data, using a stepwise regression approach to identify the most predictive of the various ALTM metrics. Separate regression models were developed for the YFE and NC study sites. These models were then applied to the ALTM data at each site to generate local maps of tree and shrub AGB.

4.2 GLAS-derived AGB

Several quality control filters were applied to the GLAS data to remove questionable data from the analysis. Because structure retrieval can be compromised by slope, waveforms from high slope (> 15°) locations were eliminated from the data pool. Waveforms returned from non-forested land cover, waveforms with a high background noise standard deviation, and waveforms fit with a single Gaussian were excluded.

A set of metrics was derived from the GLAS waveforms to use as predictor variables in a multiple linear regression to estimate AGB pre- and postfire. Canopy height was derived from the GLAS waveforms by identifying the first signal above a background noise threshold as the canopy top and the peak of the strongest return as the ground and differencing the two elevations. Canopy cover was derived by relating the energy returned from the canopy to the total energy in the waveform. The quartile metrics were calculated from the cumulative energy of the returned waveforms.

Because no GLAS footprints were coincident with the field data, extracted values from the ALTM-derived AGB maps for each coincident footprint were used as the dependent data to train the regression models. All GLAS metrics were used as the independent variables. After filtering, 403 footprints remained to build the prefire AGB model, and 658 footprints to build the postfire model. The regression models were used to calculate AGB for each GLAS footprint within the YRB area that passed the quality control filters, 6,483 for L2c and 22,549 for L3f.

A regression tree model was used to extrapolate GLAS-based tree and shrub AGB values beyond discrete footprint locations and derive AGB maps for all of the YRB study area. Regression tree analysis partitions the information space into subspaces to optimize prediction. Regression trees allow more precise modeling of complex nonlinear systems than standard multiple regressions, can provide understanding of the mechanisms that control the relationships between the independent and dependent variables, mitigate colinearity, and optimize sampling for scaling relationships from field scale to larger areas. Here, two regression trees were generated, a first using the prefire Landsat mosaic and GLAS data, and a second using postfire imagery and GLAS data. Additional predictor values that were the same for both models were elevation, slope, aspect, and the NLCD land cover data.

4.3 Pre- and postfire AGB comparison

The GLAS-derived pre- and postfire AGB data were differenced to estimate the change in AGB. The MTBS fire perimeter data were used to identify and extract subsets from the AGB difference values for each fire, and the mean of the AGB difference values was calculated. These mean values were compared to the MTBS burn severity data. The MTBS high and moderate burn severity classes were combined and the percent area within each fire represented by the combined class was calculated. The fires were then split into two groups: those where <45% of the area fell into the high-to-moderate burn severity class and those where >45% of the area fell into this class. The distributions of mean AGB difference values for the two classes were generated and an analysis of variance was performed to determine if they were statistically different. A qualitative assessment of the spatial distribution of the AGB difference values as

compared to the burn severity data was shown using the 2004 Lower Mouth Fire, located in the YFE, as an example.

5 Results

5.1 ALTM-derived AGB

The final model identified through stepwise regression analysis for estimating AGB from the ALTM data in the YFE took the following form:

$$AGB = 11 + 13*p70 - 229*d0 + 198*d1 + 256*r1 - 351*r2 + 196*r4, \quad (1)$$

where p70 was the 70th decile height, d0 and d1 are the densities for the 0th and 1st canopy layers, and r1, r2, and r4 are the canopy ratio values at 1 m, 2 m, and 4 m above ground elevation, respectively. The model had an adjusted R^2 of 0.81, and a residual standard error of 28 Mg ha^{-1} (mean = 73 Mg ha^{-1}), and was used to generate the YFE ALTM-based map of AGB. For NC, the final regression model took the following form:

$$AGB = 18.5 – 2.1 * maxht + 488 *r1, \quad (2)$$

where maxht was the maximum canopy height. The model had an adjusted R^2 of 0.81, and a residual standard error of 19 Mg ha^{-1} (mean = 26.27 Mg ha^{-1}), and was used to generate the NC ALTM-based map of AGB.

5.2 GLAS-derived AGB

The final models identified through stepwise regression analysis for estimating pre- and postfire AGB from the ALTM AGB maps and GLAS waveform metrics data took the following forms:

$$AGB_{pre} = -23 +19.7*q75 -18.5*q25, \quad (3)$$

where q75 and q25 where the 75th and 25th quartile heights of cumulative canopy energy, and,

$$AGB_{post} = -6.8 +19.4*q75 -11.5*q50, \quad (4)$$

where q50 was the 50th quartile height of cumulative canopy energy. The prefire model had an adjusted R^2 of 0.58, and a residual standard error of 24 Mg ha^{-1} (mean = 65 Mg ha^{-1}), and was used to calculate prefire AGB for the L2c footprints. The postfire model had an adjusted R^2 of 0.59, and a residual standard error of 23 Mg ha^{-1} (mean = 62 Mg ha^{-1}), and was used to calculate postfire AGB for the L3f footprints. The pre- and postfire AGB data for the Lower Mouth Fire are shown in Figure 2.

5.3 Comparison of results by severity

The distributions of mean AGB difference values for the two high-to-moderate burn severity groups are shown in Figure 3. The analysis of variance of the AGB difference values for the burn severity groups resulted in a p-value of 0.0097, indicating that they are significantly different at 95% confidence. The comparison of the AGB difference and burn severity data showed both areas of agreement and disagreement (Figure 3). An area running east/west through the center of the fire is predominantly classified as high or moderate severity and the AGB difference image shows higher change values in this region. In contrast, an area in the northeast section of the fire is classified largely as high or moderate severity, but has low AGB difference values. Most other areas with lower values of AGB difference correspond relatively well with low or moderate burn severity.

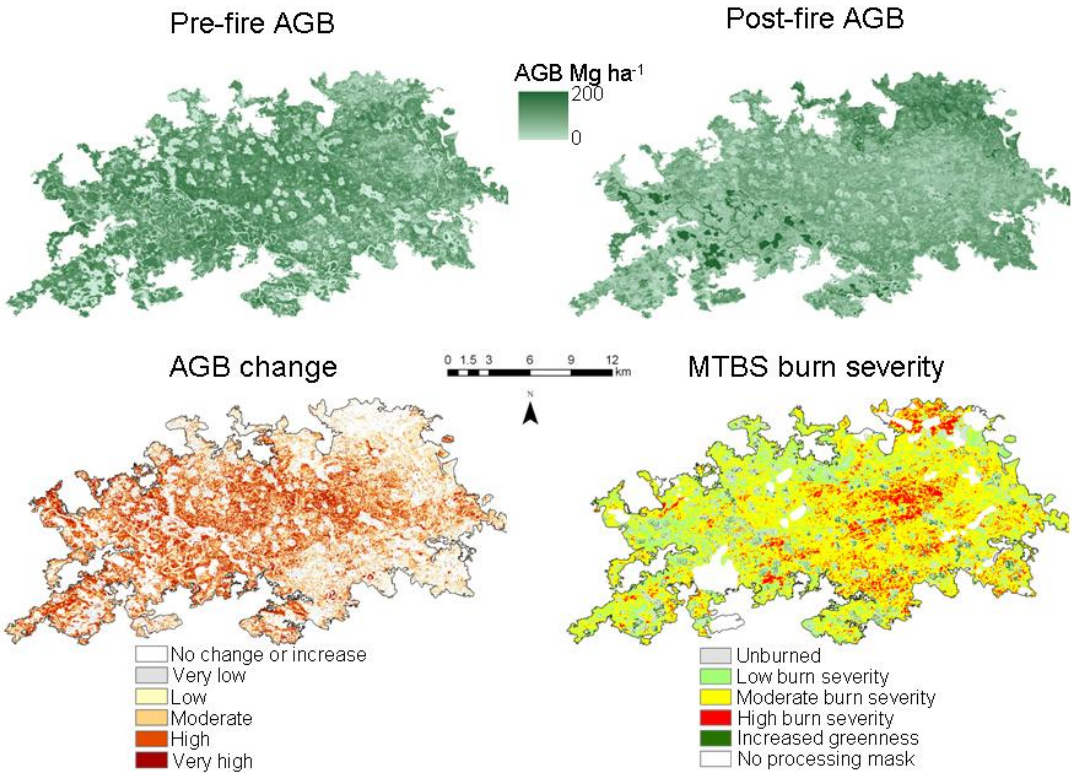


Figure 2 - Pre- and postfire AGB, AGB change, and burn severity for the 2004 Lower Mouth Fire

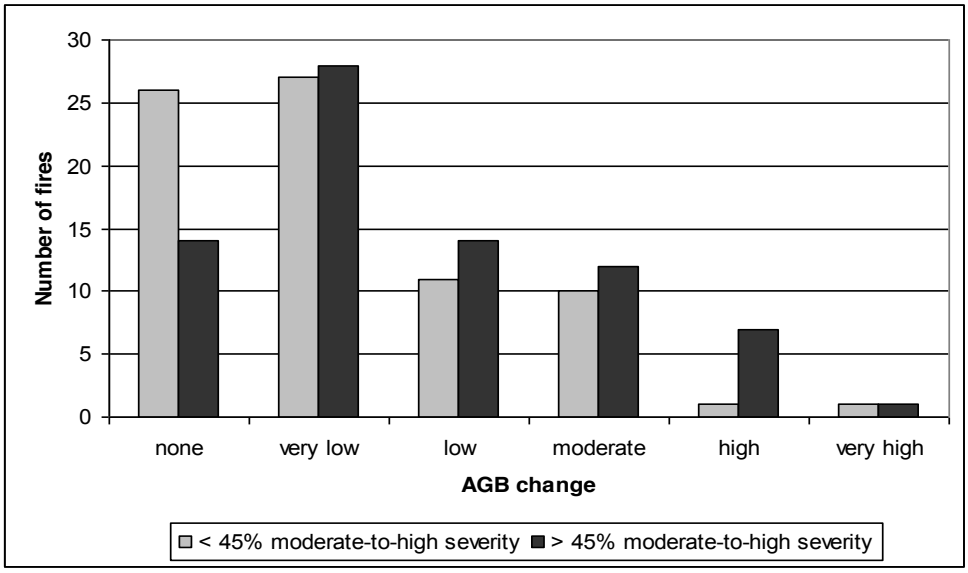


Figure 3 - Distributions of mean AGB change, per fire, separated by burn severity class

6 Discussion and Conclusion

Lidar data have been used for AGB mapping for boreal regions. For example, Næsset and Gobakken (2008) used airborne lidar data to assess above- and belowground biomass for forests in Norway, explaining 86% of the variability. Næsset *et al*.(2011) conducted AGB assessments at different landscape scales for different forest classes in Norway using lidar, resulting in R^2 values between 0.76 and 0.89. GLAS data have also been used for AGB estimation in a variety

of biomes. Lefsky *et al*.(2005) used canopy height estimates from GLAS and field observations from Tennessee, Oregon, and Brazil to predict AGB. The model explained 73% of the variability with an RMSE of 58.3 Mg ha^{-1}. Helmer *et al*.(2009) used GLAS data with other input data to assess biomass change for a tropical forest in Brazil. The observed rate of biomass accumulation matched well with those modeled. Boudreau *et al*. (2008) integrated airborne lidar and GLAS to conduct a regional biomass assessment in Québec, achieving an R^2 of 0.59.

The vegetation structure data provided by GLAS have been used for a variety of ecosystem studies and are critical to large-area vegetation structure mapping efforts, as they represent the only globally available lidar observations. However, the large footprint size of GLAS (65 m) reduces the precision of the observations compared to small-footprint, airborne systems. For example, the detail is not fine enough to provide the data necessary for fulfilling requirements for Reducing Emissions from Deforestation and Forest Degradation monitoring. Still, vegetation structure change information obtainable from GLAS, especially when used in conjunction with other remote sensing data, can provide a more realistic representation of large-scale vegetation structure patterns that those made from field observations alone.

Given current climate change predictions, wildland fire will continue to be a significant issue in Alaska and other boreal regions. Consistent, high-quality data are needed as inputs for modeling future scenarios to understand the magnitude of carbon release directly through combustion and through secondary effects such as permafrost loss and change in vegetation structure and composition. While additional work is necessary to understand underlying processes and sources of error in developing the AGB data layers, the work presented here shows how GLAS data can contribute to developing a more complete picture of the effects of wildland fire on the Alaskan landscape. In the Lower Mouth Fire example, there is considerable agreement between areas classified as high and moderate burn severity and areas identified as having a moderate to high change in AGB. These areas correspond to areas mapped as evergreen forest by NLCD. The areas of disagreement, where AGB change is low and burn severity is high, are mapped as deciduous forest. This may indicate a higher number of stems left standing postfire in the deciduous region. Such information can help identify which areas experienced crown fire and which did not, and may be an indicator of postfire recovery.

This study demonstrated how a combination of GLAS, airborne lidar and imagery can be used to assess biomass change in the YRB in response to wildland fire. To advance these research efforts, additional field work is planned for 2012 that will collect ground data coincident with GLAS footprints over a variety of vegetation and slope conditions in the YRB. These will be used to improve biomass estimation and develop a more robust accuracy assessment.

Acknowledgments

USGS Climate Effects Network and Landuse Change Research and Development Programs provided funding. Dr. Peterson's work was performed under USGS contract number G08PC91508.

References

Barrett, K., McGuire, A.D., Hoy, E.E. and Kasischke, E.S., 2011. Potential shifts in dominant forest cover in interior Alaska driven by variations in fire severity. *Ecological Applications*, 21, 2380-2396.

Bouchard, M.A., Dwyer, J.L. and Granneman, B., 2009. A Yukon River Basin Landsat Mosaic for Assessing Environmental Change. *EOS, Transactions, American Geophysical Union*, 90, Fall Meeting Supplement, Abstract GC51A-0708.

Boudreau, J., Nelson, R.F., Margolis, H.A., Beaudoin, A., Guindon, L. and Kimes, D.S., 2008. Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. *Remote Sensing of Environment*, 112, 3876-3890.

Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B. and Howard, S., 2007. A project for monitoring trends in burn severity. *Fire Ecology*, 3, 3-21.

Goetz, S.J., Sun, M., Baccini, A. and Beck, P.S.A., 2010. Synergistic use of spaceborne lidar and optical imagery for assessing forest disturbance: An Alaska case study. *Journal of Geophysical Research*, 115, G00E07.

Helmer, E.H., Lefsky, M.A. and Roberts, D.A., 2009. Biomass accumulation rates of Amazonian secondary forest and biomass of old-growth forests from Landsat time series and the Geoscience Laser Altimeter System. *Journal of Applied Remote Sensing*, 3, 033505.

Ji, L., Wylie, B.K., Nossov, D.R., Peterson, B., Waldrop, M.P., McFarland, J.W., Rover, J. and Hollingsworth, T.N., 2012. Estimating aboveground biomass in interior Alaska with Landsat data and field measurements. *International Journal of Applied Earth Observation and Geoinformation*, 18, 451-461.

Kasischke, E.S., Verbyla, D.L., Rupp, T.S., McGuire, A.D., Murphy, K.A., Jandt, R., Barnes, J.L., Hoy, E.E., Duffy, P.A., Calef, M. and Turetsky, M.R., 2010. Alaska's changing fire regime — implications for the vulnerability of its boreal forests. *Canadian Journal of Forest Research*, 40, 1313-1324.

Kasischke, E.S. and Hoy, E.E., 2012. Controls on carbon consumption during Alaskan wildland fires. *Global Change Biology*, 18, 685–699.

Lefsky, M.A., Cohen, W.B., Parker, G.G. and Harding, D.J., 2002. Lidar Remote Sensing for Ecosystem Studies. *BioScience*, 52, 19-30.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Del Bom Espirito-Santo, F., Hunter, M.O. and de Oliveira, R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, 32, L22S02.

Næsset, E. and Gobakken, T., 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, 112, 3079–3090.

Næsset, E., Gobakken, T., Solberg, S., Gregoire, T.G., Nelson, R., Ståhl, G. and Weydahl, D., 2011. Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area. *Remote Sensing of Environment*, 115, 3599-3614.

Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., Loveland, T.R., Vermote, E. and Zhang, C., 2010. Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, 114, 35-49.

Selkowitz, D.J., Green, G., Peterson, B. and Wylie, B., 2012. A multi-sensor lidar, multi-spectral and multi-angular approach for mapping canopy height in boreal forest regions. *Remote Sensing of Environment*, 121, 458-471.

Soja, A.J., Tchebakova, N.M., French, N.H.F., Flannigan, M.D., Shugart, H.H., Stocks, B.J., Sukhinin, A.I., Parfenova, E.I., Chapin, F.S. and Stackhouse, P.W., 2007. Climate-induced boreal forest change: Predictions versus current observations. *Global and Planetary Change*, 56, 274-296.

Wolken, J.M., Hollingsworth, T.N., Rupp, T.S., Chapin III, F.S., Trainor, S.F., Barrett, T.M., Sullivan, P.F., McGuire, A.D., Euskirchen, E.S., Hennon, P.E., Beever, E.A., Conn, J.S., Crone, L.K., D'Amore, D.V., Fresco, N., Hanley, T.A., Kielland, K., Kruse, J.J., Patterson, T., Schuur, E.A.G., Verbyla, D.L. and Yarie, J., 2011. Evidence and implications of recent and projected climate change in Alaska's forest ecosystems. *Ecosphere*, 2, 1-35.

Tree Height and Species Estimation Methods for Airborne Laser Scanning in a Forest Reserve

Géza Király[1], Gábor Brolly[2] & Péter Burai[3,4]

[1]University of West Hungary, Dept. of Surveying and Remote Sensing,
kiraly.geza@emk.nyme.hu
[2]University of West Hungary, Dept. of Surveying and Remote Sensing,
gbrolly@emk.nyme.hu
[3]VITUKI Environmental and Water Management Research Institute nonprofit Ltd.
[4]Károly Róbert College, Institute of Agroinformatics and Rural Development,
pburai@gmail.com

Paper Number: SL2012-148

A leaves-on and a leaves-off full-waveform airborne laser scanning campaign was carried out in a forest reserve. Different kinds of terrain and crown point-filtering, according to the full waveform information, have been tested. Several interpolation and smoothing methods have been adapted for the best Digital Surface Model for single tree-top determination. Single tree heights have been estimated when there was sufficient crown-delineation, while the heights of a group of trees have been determined otherwise. Our previous terrestrial laser scanning datasets were used to determine stem locations and leanings as the basis for tree-top estimation. The dataset produced five years ago also served as a reference concerning height growth.

The original point-cloud inside the delineated crown was analysed to determine the height. The full waveform information of these points forming a crown was investigated to estimate the tree species. The results have been compared to multispectral (WorldView-2) imagery classification. The results showed very promising basic classification possibilities on a group of species. The method was also suitable for the detection of standing dead or dying trees, which is the basis for gap-development.

The leaves-off dataset is more suitable for the single tree crown-delineation. The datasets from different phenological phases make the species determination much more accurate by factor 2.37. Further integration of the imagery data can further enhance the accuracy of the classification.

1. Introduction

New remotely sensed data have been gathered recently about a forest reserve, where several previously surveyed reference data were available. The main aim was 1) to extract information from these new data, and 2) to establish a connection between this new information and the previous, mostly terrestrial data.

There have been a number of publications on individual tree detection for the past decade. Kaartinen et al (2012) provided a very good overview and a comparison of the existing methods and their accuracy. However, we have not found any publications on the linking of the results of tree detection and field references in the case of leaning trees. This is the reason why we have tried to further develop our previous method (Király et al 2007) with the aim of linking our results to the terrestrial reference data.

Locating trees from ALS data is future topic for research on forest stem dynamics as natural processes can be evaluated on individual trees over large areas. Kaartinen et al (2012) provide an extensive overview on existing algorithms aiming at stem detection from ALS data. However; the tree top detection is still challenging and further studies are needed to find algorithms that can exploit the full advantages of detailed surface models from high density laser scans.

The phenologic state of deciduous and mixed stands strongly affects the quality of DTM, DSM and additional information to be extracted from ALS data. Through the combination of full-waveform scans recorded in the leaf-off and leaf-on state, phenologic characteristics of trees can be considered. This additional information can be used for the classification of species groups and to enhance the delineation of individual trees.

An individual tree map with corresponding height estimates were created based on TLS data in our previous research studies (Király & Brolly 2007, Brolly & Király, 2009), where tree positions were located at breast height (1.3 meters aboveground). In order to build the interconnection between the tree positions from the existing data base and the tree tops resulting from the evaluation of ALS data, the corresponding locations are to be matched with regard to the local DSM morphology and the leaning of the stems.

2. Methods

2.1 Study area

The Hungarian Forest Reserve Program and Network was initiated by NGOs, researchers and ministerial stakeholders in the 1990s. The network consists of 63 forest reserves, which represent the most characteristic (semi-natural) forest types of Hungary and the Pannonian region, to protect and study natural or semi-natural forest ecosystems (See Figure 1; Horváth et al. 2001).

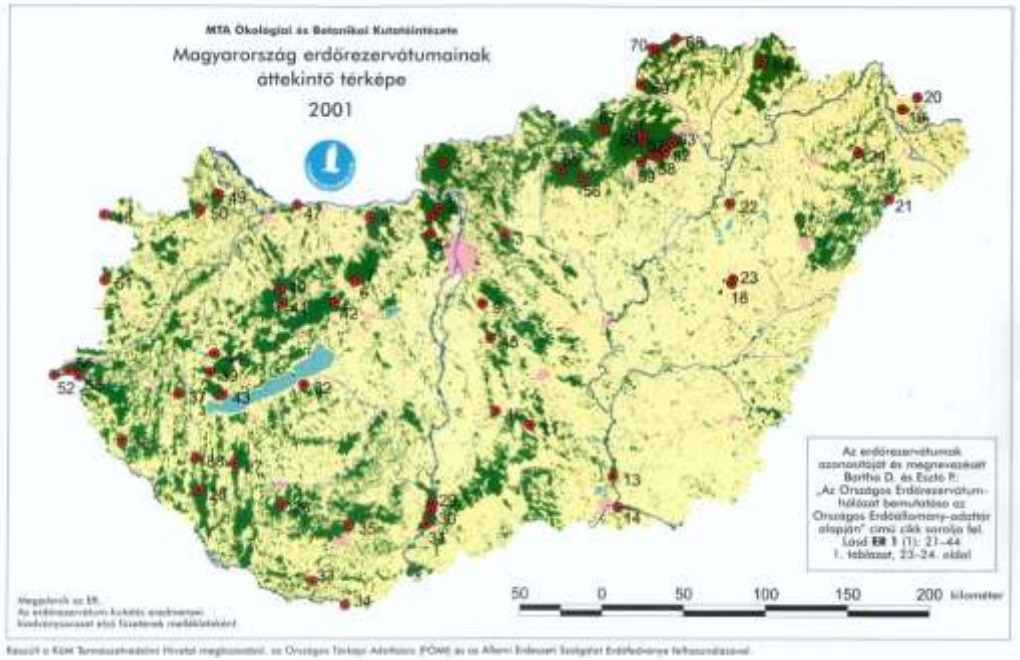


Figure 1: The overview map of the 63 Forest Reserves in Hungary
(http://www.erdorezervatum.hu/kepek/attekinto_1000.JPG)

Some of these reserves are selected for long-term investigations, and a permanent sampling network has been established. The Forest Reserve No 46, called 'Hidegvíz-völgy' (Cold-creek valley), situated in the North-West corner of Hungary, right on the border of Austria is among those selected. This reserve has a small, 19.7 ha core area, and a small, 37.2 ha buffer area; resulting in a low ratio of 1.89 of buffer/core (1.89) where 3 is the suggested value (Bartha and Esztók, 2001), but one side of the area is exposed because of the Austrian border (see also Figure 3). Abandoned old beech stands are the characteristic type in the Sopron Mountains with about 750 mm annual precipitation and 8.5 ºC mean annual temperature. Some introduced spruce stands are also present, but perishing in patches due to bark beetle disease. Lessivated, acidic, non-podzolic brown forest soils with surface water gley, and colluvial soils formed on gneiss and mica bedrock typically occur on a side slope of a brook flowing South- or South-East-ward (Horváth et al. 2012) (see Figure 3).

2.2 Reference data

A regular, 50 * 50 m grid of permanent sampling points (PSP) were planned and set up in the selected forest reserve. Our department made the PSP establishment in this No. 46 Reserve (Király 2006) in order to gain information about the pattern and dynamics of the tree populations and stand structure. Four modules of standard thematic survey were interlinked in these PSPs: stand structure (including standing and lying dead wood), sampling of the shrub and regeneration layer, sampling of ground vegetation and site evaluation. We plan to re-survey these PSPs in 10–15 years to follow up and evaluate natural changes in these forest stands. This approach is called FOREST+n+e+t – the stand dynamic and ecological observation network of natural forests.

A terrestrial laser scanning campaign was done in 2006, and a couple of articles have been published on the topic (e.g. Király et al 2007, Király and Brolly 2007, Brolly and Király 2009). A detailed single-tree based database has been created from field surveys and terrestrial laser scanning. There are 216 single trees in the reference data around one PSP, and 564 single trees with fewer attribute data around another PSP (see Figure 3).

2.3 Remote Sensing data

Recently the study area has been surveyed with Airborne Laser Scanning on two occasions in the frame of the ChangeHabitats2 project (http://www.changehabitats.eu/). The parameters of the surveys can be found in Table 1.

Table 1: The parameters of the Airborne Laser Scanning Surveys

Parameter	Leaves-on	Leaves-off
Date	2011.07.12	2012.03.26
HAG	~660 m	~630 m
Overlap	~44%	~54%
Ave. point density	22.3 points/m^2	35.7 points/m^2
Sensor	Riegl LMS-Q680	Riegl LMS-Q680

A more detailed picture is provided about the point density and overlaps in Figure 2

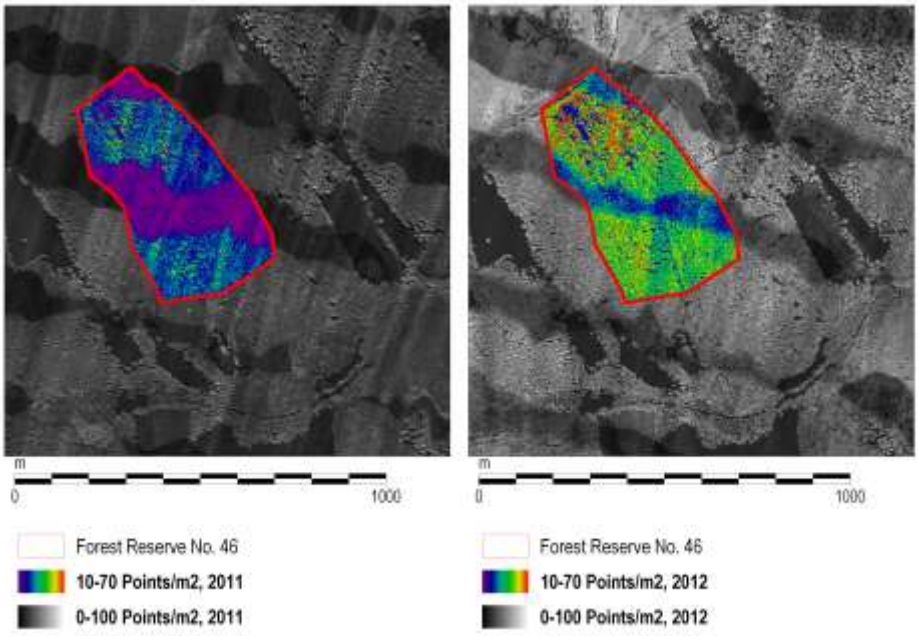


Figure 2: The point densities of the summer (left) and winter (right) flight

Two WorldView-2 imageries were collected additionally on the 17th of August 2011; and a hyperspectral flight was carried out on the 26th of August, 2011. The images were collected in the frame of the R+D project (TÁMOP - 4.2.1.B-09/1/KONV).

2.4 Creation of the Digital Elevation Models

The base dataset for creating the Digital Elevation Model was the year 2012 leaves-off, a bit denser scanning data.

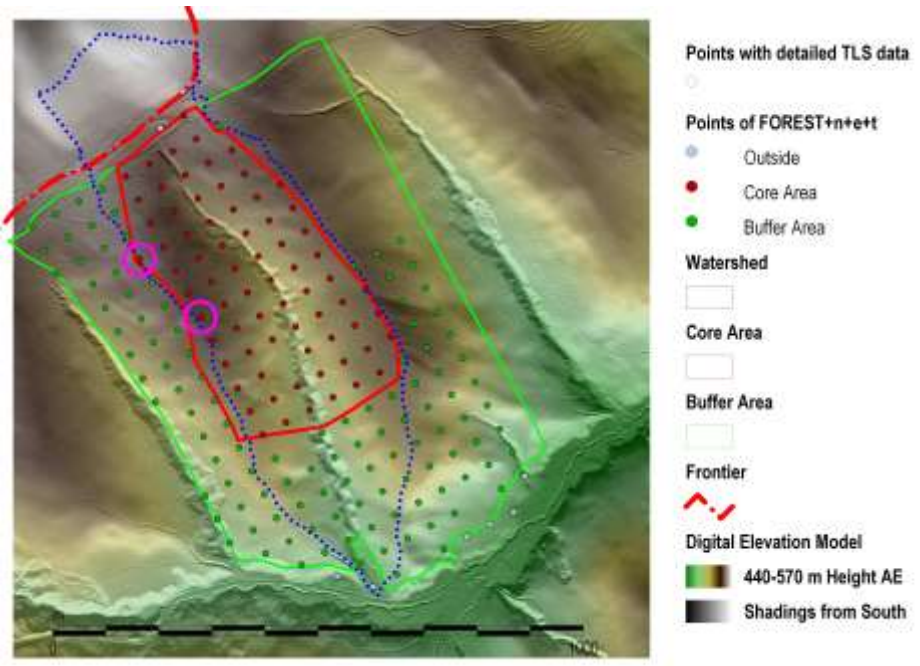


Figure 3: The Digital Elevation Model, the Reserve's border and the PSPs of the area

Firstly a filtering of the input points was applied, selecting the last echos, which are narrower than 4.5 ns echo-width. Then the hierarchic robust interpolation of the SCOP++ software (Kraus and Pfeifer 1998; Briese et al 2002) was used, with 0.5 m grid resolution. The result is shown in Figure 3.

The watershed was calculated based on the new digital elevation model (see blue dotted border in Figure 3), because, theoretically it should coincide with the frontier of Austria and Hungary on the NW.

2.5 Creation of the Digital Surface Models

Different interpolation and filtering methods have been applied for the Digital Surface Model (DSM). The following DSM interpolation methods were applied:

1. (Snap)
2. Nearest Neighbours
3. Moving average (horizontal plane)
4. Moving planes (tilted plane)
5. Matching paraboloide
6. SCOP++ LIDAR DSM
7. Spline

We started with the interpolation methods as different options in OPALS (Mandlburger et al 2009). 1) A snap grid is the simplest method. For each ALS point the appropriate grid cell is determined and the height of the ALS point is mapped to this cell. 2) Nearest Neighbour – for each grid cell the nearest ALS data point is queried from the ODM and the height of this nearest neighbour point is mapped to the grid cell. 3) Moving average – for each grid cell *n* nearest ALS points are queried and the average height of all *n* points is mapped to the grid cell. 4) Moving planes – for each grid cell the *n* nearest ALS points are queried and the best fitting tilted plane (minimizing the vertical distances) is estimated. The height of the resulting plane at the grid point (x,y) position is mapped to the grid cell (OPALS Manual). 5) Moving paraboloide – we could not get a correct result from OPALS (using v1.1) with this interpolation, so we created our own matching paraboloide interpolation.

Our ‘matching paraboloide’ method works as follows. For each grid cell we find the points within the search radius, then calculate the statistics of the selected points, remove the points below the mean-1σ and the highest point. After that filtering, we check if there are enough points in all four quadrants. If yes, we fit the paraboloide (equation (1) and (2)), if no, we fit a plane to avoid gaps in the interpolated model.

$$f(x,y) = c_0 + c_1 \cdot x + c_2 \cdot y + c_3 \cdot x \cdot y + c_4 \cdot x^2 + c_5 \cdot y^2 \quad (1)$$

Where: $c_{0\text{-}5}$ are the coefficients of the second order surface

$$\sum_{j=0}^{n} (z_j - f(x_j, y_j))^2 \quad (2)$$

Minimizing the vertical deviations of the points from the fitted paraboloide results in a linear equation with six unknown elements.

There is an interpolation method in SCOP++ software named LIDAR DSM, which produces a surface model with just one level of robust filtering (Kraus and Pfeifer 1998). The last method was a Spline interpolation. After a very simple point filtering, where the second highest points

from all first echoes had been selected for each grid cell, a regularised spline interpolation was applied, creating a minimum curvature smooth surface (with a smooth first-derivative surface, too), which passes exactly through the data points (Franke 1982).

2.6 Segmentation of the crowns

The normalised Digital Surface Models (nDSM or Canopy Height Model - CHM) were created by the simple subtraction of the models. Most of the calculations were performed at 0.5 m resolution. The inverse of the normalised Digital Surface Model was necessary for modelling the watershed, so the following equation (3) was used:

$$\mathrm{nDSMneg} = \mathrm{DEM} - \mathrm{DSM} \tag{3}$$

Where: nDSMneg is the inverse nDSM (nDSM negate)

An inverse watershed modelling was applied to delineate individual tree crowns (Gougeon 1995, Király 1998). In these works were based on very high resolution aerial images with adequate lighting conditions. Treating the image as an inverse digital terrain model, where the brightest pixels are the lowest points, the individual tree crowns are the watersheds, which can easily be separated by well-known algorithms (Jenson and Domingue 1988). These image-based methods were adapted successfully with airborne laser scanning data, where the tops of the trees were determined with the Digital Surface Model (Hyyppä et al. (1999)).

The sink point of a watershed was considered as the tip of the tree and the watershed itself was regarded as the crown of the tree (see Figure 4).

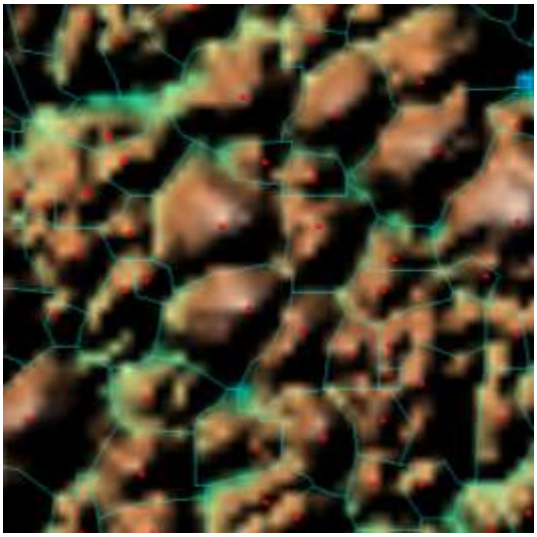


Figure 4: Shaded detail of an nDSM (winter) and the tree tips (red) and crowns (cyan) of interpolation ‘Matching paraboloide’

2.7 Tree height estimation

The co-registering of the tree tips and crowns with the reference data proved to be a very important, but complicated step. Our previous investigations had shown that the leaning of the tree makes this step complicated (Király and Brolly 2007). For this reason we utilised the modelled trunk and its leaning as a simple function of the height (Brolly and Király 2009), and found the nearest tree tips to the projected tree top. There are several ways to determine the

heights, and the most common ways utilising nDSM are as follows:

Starting from the root (terrestrial reference)

1. The tree tip is right above the root
2. Find the nearest (from horizontal point of view) tree tip (sink-point) to the root
3. Project the axis of the modelled trunk to the DSM and then find the nearest tree tip (sink-point) to that intersection

Starting from the tip (airborne reference)

4. Determine the height right from the DSM; starting from the tree tip (sink-point), and find the nearest terrestrial reference tree
5. Determine the height right from the DSM; starting from the tree tip (sink-point), the correct terrestrial reference is already known

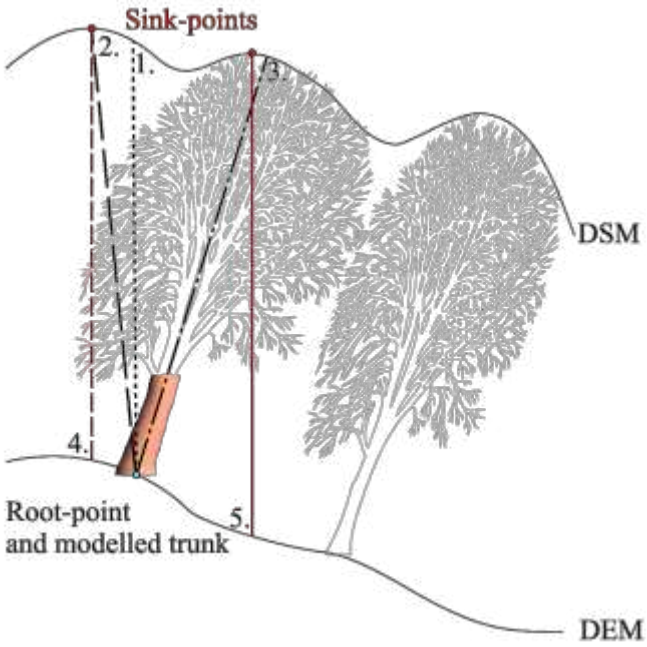


Figure 5: Comparison of the different height calculation methods (revised from Király et al 2007.)

Please note, that there can be significant differences between these heights, even on undulating area. No. 2 and No. 3 together with No. 4 and No. 5 can be the same, supposing the nearest sink-point is identical to the one determined by the modelled trunk. This is common with approximately 2/3 of the trees.

2.8 Tree species estimation

The four main tree species in the reserve, which are more or less in dominant or co-dominant situations, are:

1. European Beach (*Fagus sylvatica*)
2. Sessile Oak (*Quercus petraea*)
3. Norway Spruce (*Picea abies*)
4. European Larch (*Larix decidua*)

Certain crowns have been selected from each species as a train dataset for the full waveform (FWF) classifications. The Amplitude (Amp) and the Echo Width (EW) were used from the available FWF information of the original point cloud. Regular grid-based representations of these attributes (Amp and EW) have also been created.

The Echo Width does not show any significant difference between these main tree species, but leaves-off echoes are generally a little bit narrower (e.g. 2.19 and 2.07 for beach). The Amplitude of the training crowns is shown in Figure 6.

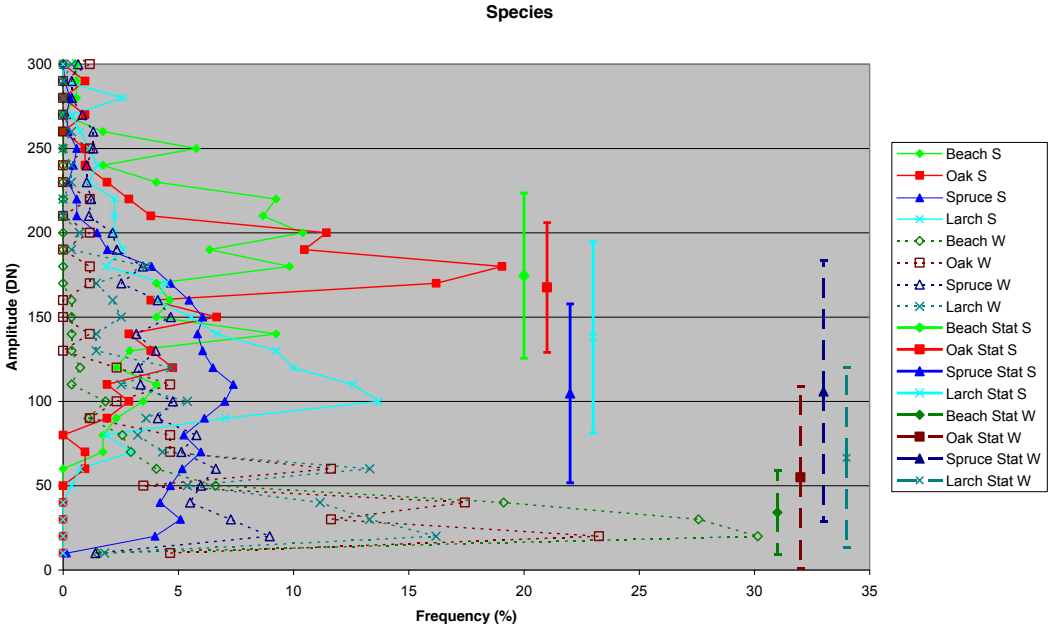


Figure 6: The Amplitude distributions and statistics for the main dominant tree species in summer (S) and winter (W) first echoes

The followings can be stated clearly:

- The beach and oak (deciduous) trees are very close to each other, both in summer- and wintertime;
- The spruce has almost identical mean values in summer and winter, but the standard deviation is much higher in winter;
- The mean values of larch occur between the deciduous and the spruce, both in summer and winter.

A dual thresholding algorithm was applied on the mean amplitude values which were calculated for the segmented crowns based on summer and winter points.

3. Results and discussion

The results are presented in three consecutive subsections. The first one is concerned with crown detection, which is the basis for the other two, the height and the species estimations.

3.1 Crown detection

The success of tree tip and crown detection basically depends on the quality of DSM. Generally the more developed interpolation gives better results. The first interpolation provided poor results, however, this was not a real interpolation. The 2nd and 3rd gave moderate accuracy. We achieved good results with the 4th one, the moving planes. The last three produced outstanding detection accuracy. The results are in the following table, where the successful detections of conifers are shown (see Table 2)

Table 2: The results of crown detection

No	Methods	Leaves-on	Leaves-off
1	Snap	53.5%	69.8%
2	Nearest Neighbour	83.7%	79.1%
3	Moving average	83.7%	83.7%
4	Moving planes	86.0%	86.0%
5	Matching paraboloide	**93.0%**	**95.3%**
6	SCOP++	**95.3%**	90.7%
7	Spline	93.0%	90.7%

It should be noted that the differences are small. The SCOP++ interpolation produced the best result on the summer points, while the 'Matching paraboloide' gave square results from both datasets. The shapes of the crowns correspond in No. 5 and No. 6, but sometimes irregular in No. 7.

3.2 Height estimations

The statistical results of the height differences from the point-based approach on the winter dataset (which was treated as a reference) determined by different methods are shown in the following table (see Table 3.)

Table 3: The comparison of the height estimations

	S-W	M1	M2	M3	M4	M5	2006
Min	-0.966	-10.568	-4.076	-1.427	-3.973	-1.307	-1.974
Max	1.387	-0.166	1.913	0.481	1.855	0.484	0.939
Mean	**-0.017**	-2.565	-0.639	**-0.513**	-0.577	**-0.425**	**-0.591**
Std	0.335	1.716	0.870	**0.312**	0.878	0.332	0.455

The point-based summer and winter results are very close to each other; their mean deviation is just 1.7 cm. Theoretically the 3rd height-calculation method (M3) is the correct one, however the 5th one (M5) produced slightly better results. The M4 is slightly better then M2 among those without terrestrial reference data. These generally occur due to higher variations in nDSM are much higher than in DEM in this area. Nevertheless, there are also some remarkable deviations in certain cases (see Figure 7).

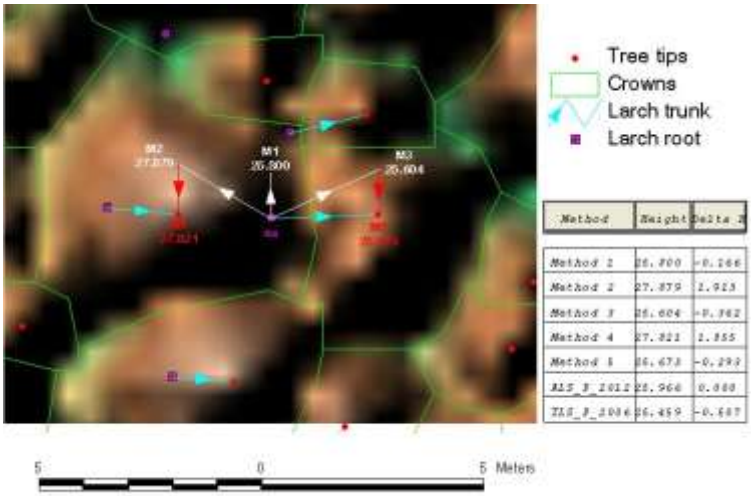


Figure 7: The differences between the height estimation methods

There are almost 2 metres of difference in the nearest neighbour methods (No. 2 and No. 4). Consequently, the link between the terrestrial and airborne data appears crucial. The average vertical growth of the sample trees is about 60 cm in six years.

3.3 Species estimations

The result of the crown species classification is shown in figures (see Figure 8 and Table 3)

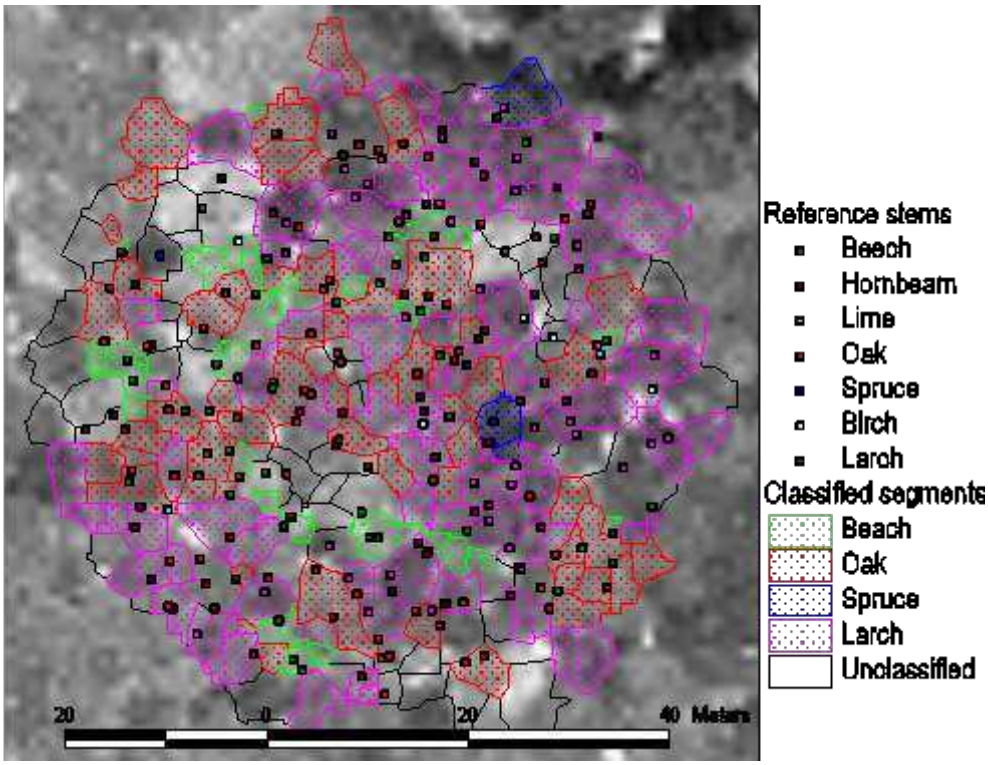


Figure 8: The classified crown segments of the sample area on the rasterized summer amplitude data

Table 3: The results of species classification

		Reference				
		Beach	Oak	Spruce	Larch	Sum
Classes	Beach	25.65%	0.00%	0.00%	0.00%	4.84%
	Oak	74.35%	100.00%	0.00%	1.14%	18.65%
	Spruce	0.00%	0.00%	100.00%	0.00%	2.35%
	Larch	0.00%	0.00%	0.00%	98.86%	74.16%
	Sum	18.86%	3.77%	2.35%	75.02%	100.00%

It is obvious from both figures that separation of the deciduous species based on the Amplitude data is not possible, but it can easily be improved with the high spectral resolution WorldView-2 imagery.

Acknowledgements

Many thanks to the ChangeHabitats2 Project for offering the ALS data about the Sopron Mountains, and the TÁMOP - 4.2.1.B-09/1/KONV project for supplying the imagery data.

References

Bartha, D. and Esztó, P., 2001. Az Országos Erdőrezervátum-hálózat bemutatása az Országos Erdőállomány-adattár alapján (Presentation of the Forest Reserve Network Hungary based on the Hungarian Forest Resource Database). In: *Az erdőrezervátum-kutatás eredmáényei (Results of the Forest Reserve Research)* ER 1(1) MTA ÖBKI, Vácrátót, Hungary, 21-44

Briese, C., Pfeifer, N., and Dorninger, P. 2002. Applications of the Robust Interpolation for DTM determination, in: *International Archives of Photogrammetry and Remote Sensing, Photogrammetric Computer Vision (PCV'02)*, 9–13 September 2002, Graz, Austria,

Brolly, G., and Király, G., 2009. Algorithms for Stem Mapping by Means of Terrestrial Laser Scanning. In *Acta Silvatica Et Lignaria Hungarica* Vol 5. Sopron, 2009. pp 119-130

Franke, R., 1982. Smooth Interpolation of Scattered Data by Local Thin Plate Splines. In: *Comp. & Maths. with Applls.* 8. (4): 237 – 281

Horváth, F., Mázsa, K., Temesi, G., 2001. Az erdőrezervtáum-program (Forest Reserves Programme in Hungary). In: *Az erdőrezervátum-kutatás eredmáényei (Results of the Forest Reserve Research)* ER 1(1), MTA ÖBKI, Vácrátót, Hungary. 6-20.

Horváth, F., Bidló, A., Heil, B., Király, G., Kovács, G., Mányoki, G., Mázsa, K., Tanács, E., Veperdi, G. and Bölöni, J., 2012. Abandonment status and long-term monitoring of strict forest reserves in the Pannonian biogeographical region. In: *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*: Official Journal of the Societa Botanica Italiana, DOI:10.1080/11263504.2011.650728

Hyyppä, J., and Inkinen, M., 1999. Detecting and estimating attributes for single trees using laser scanner. *The Photogrammetric Journal of Finland*. 16 (2): 27-43.

Jenson, S. K., and Domingue, J. O., 1988. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis. In: *Photogrammetric Engineering and Remote Sensing*. 54 (11): 1593-1600

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, Ch., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B. M., Wu, J-C., 2012. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. In: *Remote Sens.* 2012, 4, 950-974; doi:10.3390/rs4040950

Király, G., 1998. Multiscale images in forestry. In: *ISPRS International Archieves of Photogrammetry and remote sensing*. XXXII, 7. ISPRS Comission VII Symposium on „Resource and Environmental monitoring." Budapest, Hungary. 1-4 September, 1998. 365-369

Király, G., 2006. ERDŐ+h+á+l+ó létesítése a Hidegvíz-völgy Erdőrezervátum magterületén és a védőzóna kiválasztott területein (Establishing the FOREST+n+e+t in the Cald-creek Valley). MTA ÖBKI. Research report. p. 3 + 1 map

Király, G., Brolly, G. and Márkus, I., 2007. Földi lézerszkennelés alkalmazása egyes fák vizsgálatára (Application of Terrestrial Laser Scanning for Investigation on Single Trees). In *Geomatikai Közlemények X (Bulletins of Geomatics, X.).* MTA GGKI, Sopron, 2007. pp. 241-250.

Király, G. and Brolly, G., 2007. Tree height estimation methods for terrestrial laser scanning in a forest reserve. In: "*Proceedings of ISPRS Workshop on laser scanning 2007 and SilviLaser.*" Espoo, Finland. 12-14 September, 2007. 211-215.

Koch, B., Heyder, , U. and Weinacker, H., 2006. Detection of Individual Tree Crowns in Airborne Lidar Data. In: *Photogrammetric Engineering and Remote Sensing*, 72 (4), 357-363.

Kraus, K., and Pfeifer, N., 1998. Determination of terrain models in wooded areas with aerial laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* 53, pp. 193-203.

Mandlburger, G., Otepka, J., Karel, W., Wagner. W. and Pfeifer, N., 2009. Orientation and Processing of Airborne Laser Scanning data (OPALS) - concept and first results of a comprehensive ALS software. *ISPRS Workshop Laserscanning* 2009, Paris, France.

OPALS Manual: OPALS - Orientation and Processing of Airborne Laser Scanning data http://www.ipf.tuwien.ac.at/opals/html/index.html

Assessing the impact of broadleaf tree structure on airborne full-waveform small-footprint LiDAR signals

Paul Romanczyk[1], David Kelbe[1], Jan van Aardt[1], Kerry Cawse-Nicholson[1], Joe McGlinchy[2], and Keith Krause[3]

[1]Chester F. Carlson Center For Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY, USA, 14623 par4249@rit.edu, djk2312@cis.rit.edu, kacpci@cis.rit.edu, vanaardt@cis.rit.edu
[2]ESRI jmcglinchy@esri.com
[3]National Ecological Observatory Network kkrause@neoninc.org

Paper Number: SL2012-149

Abstract

Airborne small-footprint waveform light detection and ranging (LiDAR) research in forest environments is still at an early stage of development, mainly because the complex signal-matter interactions are not fully understood. This lack of understanding is due to the high temporal resolutions of scan interactions, off-nadir scanning, and attenuation of the signal energy as it propagates through a canopy. However, with the increased use of small-footprint waveform LiDAR systems for forest assessment, it is important to understand and quantify how the tree structural components impact the received backscattered signal. Quantifying the impact of these structural components on the waveform will define the limit of what can be detected from a waveform sensor, e.g., leaf area index (LAI), branch structure detail, canopy parameters, etc. We therefore present a simulation study that evaluates the effect of structural components on the received signal from an *Acer rubrum* (red maple) and *Quercus rubra* (red oak) stand.

It was found that the impact of leaf stems and twigs is not statistically significant at a 95% confidence level. The other components (trunks, boughs, branch-level-1, branch-level-2, and branch-level-3) do have a statistically significant contribution to a waveform LiDAR signal. The leaves represent a significant component of the backscattered waveform signal. Finally, increasing the outgoing pulse width decreases the differences between observed waveforms for various levels of tree component geometry. These results will help to better define the construction of virtual forest scenes in simulation studies, while also contributing to an improved understanding of LiDAR-forest interactions.

1. Introduction

In recent years, there has been an increase in the use of airborne Light Detection and Ranging (LiDAR) for acquiring remotely sensed data. Small-footprint full-waveform LiDAR is a relatively new technology that is only beginning to be more commonly used due to its high temporal resolution of the backscattered laser pulse (Mallet and Bretar 2009). Despite the increased use of small-footprint full-waveform systems, the types of interactions that affect the backscattered signal are not completely understood.

A few studies such as Sun and Ranson 2000 and Ni-Meister *et al*. 2001 have looked into the effect of geometry on large-footprint (10-25m) full-waveform LiDAR signals. Sun and Ranson 2000 found that stand structure and higher order scattering affect the waveform. Ni-Meister *et al*. 2001 found that leaf orientation and clumping contribute to the waveform signal. In this paper

we propose a method that evaluates the effect of various plant components on small-footprint (0.5m), full-waveform LiDAR signals.

Our objectives were to (i) assess the level of structural detail, i.e., leaf vs. bole vs. various orders of branching, that can be resolved by typical waveform LiDAR sensors and (ii) evaluate the impact of outgoing pulse width, signal attenuation, and noise on quantification of such structural parameters. We hypothesized that backscatter will be dominated by foliage, but that large branches will also impact the signal; this latter interaction is expected to be a function of range and outgoing pulse width. However, it is practically infeasible to sequentially remove tree structural components in order to perform such a signal-matter assessment. We therefore made use of simulated waveforms using the Digital Imaging and Remote Sensing Image Generation (DIRSIG) simulation environment to generate synthetic waveforms in the initial phase of this study. The virtual scene is based on forest inventory data from Harvard Forest, MA, USA. Species include *Acer rubrum* (red maple) and *Quercus rubra* (red oak).

2. Methods

2.1 Scene Construction

The goal of the virtual scene is to have it be representative of a forest stand that could be found in nature. To accomplish this, we based our trees on the height, diameter at breast height (DBH), and species from a Harvard Forest inventory (Munger and Wofsy 1999). We selected nine *Acer rubrum* (red maple) and four *Quercus rubra* (red oak) trees to model. The parameters from the Harvard Forest inventory were used as a basis for creating the tree models in OnyxTREE (Bosanac and Zanchi 2011). The tree geometry was exported on a per-component basis (trunk, boughs, branch-level-1, branch-level-2, branch-level-3, twigs, leaf stems, and leaves). A one character identifier is used for each of the components: trunk (t), boughs (b), branch-level-1 (1), branch-level-2 (2), branch-level-3 (3), twigs (w), leaf stems (s), and leaves (l). The trees were manually placed in the scene so that the canopy is mostly closed. Appropriate reflectance and transmission spectra, sampled from an existing spectral library were attached to the respective components. The tree components were placed on an infinite ground plane with an associated grass spectra. The scene is circular with an approximate radius of 25m. Figure 1 shows a simulated RGB image of the scene.

Different combinations of geometries were used for the simulations: all geometry (tb123wsl); no leaf stems (tb123wl); no leaf stems or twigs (tb123l); no leaf stems, twigs, or branch-level-3 (tb12l); trunks, boughs, branch-level-1, and leaves (tb11l); trunks, boughs, and leaves (tbl); and leaves only (l).

2.2 Simulations

2.2.1 DIRSIG

The Digital Imaging and Remote Sensing Image Generation (DIRSIG) model is a physics-based ray-tracing program used for producing spatially, spectrally, radiometrically, and temporally correct synthetic images (Schott *et al*. 1999). The DIRSIG waveform LiDAR simulation employs a two-pass hybrid of forward and backward Monte-Carlo ray tracing (Brown *et al*. 2005).

For the forward part of the simulation, a photon is randomly cast into the scene. The photon will then follow a random walk based on the properties of the material(s) that it encounters (transmission, reflection, or absorption). An event in a 3D data structure is created for each interaction of the photon with the scene. The photon progresses through the scene until it is

absorbed or the user-defined maximum number of interactions is reached. The laser transmitter model defines the spatial, spectral, and temporal distribution of the photons.

The second pass of the LiDAR simulation is a backward propagation from each detector element in the sensor. A ray is projected from the detector, through the optics until it encounters geometry in the photon-mapped scene. The photon map at the location of intersection is queried, and the number of photons, as well as the time of the interaction is added to the signal. It is possible to use subsampling to have better sub-pixel spatial or temporal fidelity. In addition to the active (LiDAR) terms, the passive (sun, moon, skylight) values are also collected. The waveform produced by DIRSIG is an estimate of the mean photon arrival at the sensor. The signal does not take into account any electronic noise.

2.2.2 DIRSIG settings

The scene was placed at 43.531911° N, 72.190711° W. The elevation of the scene is 0m above the geoid. A mid-latitude summer (mls) was used for the weather conditions. A simple atmospheric model (no MODTRAN) with a 270K apparent sky temperature was used for the radiation transport.

The LiDAR sensor was placed at 1000m at nadir over each site. The outgoing laser pulse was Gaussian with a mean of 1064nm and spectral line width of 0.01nm. The pulse energy was 0.0002J. Outgoing pulse widths of 4ns, 8ns, and 16ns were used for this assesment. The spatial profile of the beam was radially uniform with a divergence half angle of 0.00025rad. This corresponds to a 0.5m spot size on the ground. A single pixel detector with 100x100 subsampling within a 0.5m ground sample distance (GSD) was used. The sensor had the time gate open after 6.404e-06s and closed after 6.738e-06s. A sample was taken every 1ns. This corresponds to a distance range of about 960m to 1010m, sampled every 15cm for a total of 335 time bins. The simulated collect time was 12:00 PM UTC-5 (US Eastern) on June 1, 2012. There were 250,000 maximum source bundles, with a maximum of four bounces per bundle. Photon mapping was used for higher order returns only. There were a maximum of 500,000 events in the map.

For each site and outgoing pulse width combination, 30 scans were taken of the validation geometry (tb123wsl; all tree structure components included) to estimate photon arrival statistics. An additional scan for each site/outgoing pulse width combination was acquired for each reduced geometry set in order to compare the effect of the backscatter on different geometry levels to the validation signal.

2.2.3 Sampling

In order to maintain reasonable computation times, ten locations were randomly selected from the scene. Five of the sites were drawn from a radially symmetric uniform distribution with radius 15m and the other five from a radially symmetric uniform distribution of radius 25m. The goal of the sampling was to have the majority of the points lying within the canopy. Eight of the sites lie on the canopy. Sites 1 and 4 are in a small gap within the canopy. Site 5 is outside the canopy, but inside of the shadow of the canopy. See Figure 1 for the distribution of the sites within the scene.

2.3 Comparison Metrics

2.3.1 Spectral Angle Mapper

The mean of the “all” geometry (tb123wsl) was used as the truth for each site/outgoing pulse

width combination. Spectral Angle Mapper (SAM) is a method to find the angle between two vectors in multidimensional space (Richards and Jia 2006). A SAM value close to 0° indicates a similarity between waveforms. A SAM value close to 90° indicates an absolute dissimilarity between waveforms.

$$\mathrm{SAM}(\mathbf{x},\mathbf{y}) = \cos^{-1}\left(\frac{\mathbf{x}\cdot\mathbf{y}}{||\mathbf{x}||\,||\mathbf{y}||}\right) \tag{1}$$

where **x** and **y** are vectors, and ||**x**|| is the norm of **x**. In this case, **x** is the mean waveform of the validation set and **y** is the waveform of a different set of geometries.

Figure 1: A simulated RGB image of the Harvard Forest scene. Locations of random sites are shown. The circles represent the size of a single pulse on the ground.

2.3.2 Prediction Intervals

A prediction interval was used in order to have an understanding of where the waveforms differ from the validation signal. The bounds of the prediction interval from the mean are given by

$$\mathrm{P.I.}(\alpha, m, s, n) = m \pm t_{\alpha/2,n-1}\cdot s\cdot\sqrt{1+\frac{1}{n}}, \tag{2}$$

where m is the sample mean, t is the critical value from the student's t distribution with (1-α)×100% confidence levels, n is the number of samples, and s is the sample standard deviation. A value is not from a standard normal with mean m and standard deviation s at a (1-α)×100% confidence level if it lies outside of the prediction interval.

Since the DIRSIG simulated waveforms represent mean photon arrivals, the standard deviation of the validation signals is small (the waveforms are consistent when using the same parameters). Assuming that photon arrival is Poisson distributed (Schott 2007), the shot noise

component of the standard deviation of the waveform will be the square root of the mean waveform. The shot noise component of the standard deviation is large compared to the variation between simulations; therefore the combined standard deviation will approximately be the standard deviation of the shot noise.

3. Results

Figure 2 shows representative simulated waveform signals from our scene. There are two cases: one which has a small amount of tree structure (site 1, (a)) and another which has a large amount (site 2, (b)). In all geometry cases, the shape of the simulated waveform is approximately the same even though there are different amplitudes for the peaks. As the amount of geometry decreases, the energy in the waveform shifts towards the ground, due to the increased probability of photon penetration through canopy.

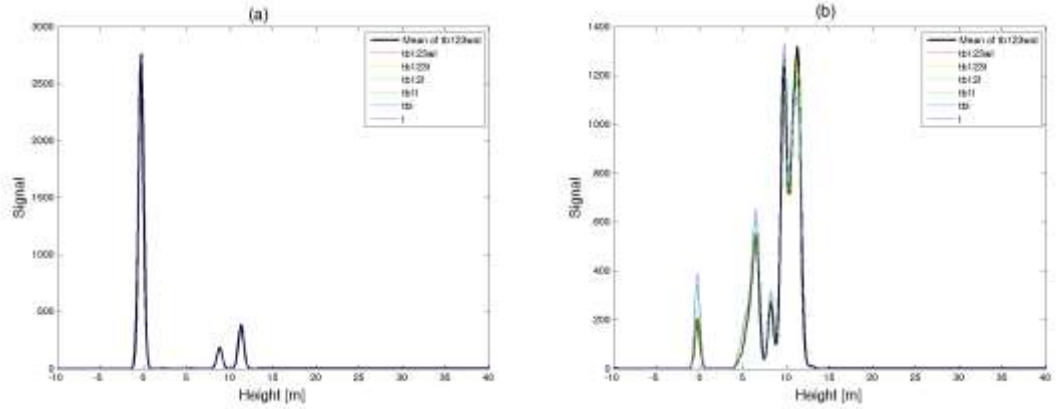


Figure 2: Simulated waveforms for various levels of geometry where (a) is site 1, (b) is site 2. Both are 4ns outgoing pulse width. E.g., note the differences in simulated waveforms for different levels of geometry.

The difference between the mean of the validation set and the geometry with twigs and leaf stems removed is shown in Figure 3. It was found that the tb123wl and tb123l geometries are not statistically different from the validation waveform across all sites and pulse widths. Figure 4 shows the differences between the mean of the validation set and leaf only geometry for the same two sites. For site 1 (in a gap in the canopy), the leaf-only geometry was not statistically different from the validation set mean (Figure 3a). The site 2 leaf only-geometry was statistically different at a 95% confidence level (Figure 3b).

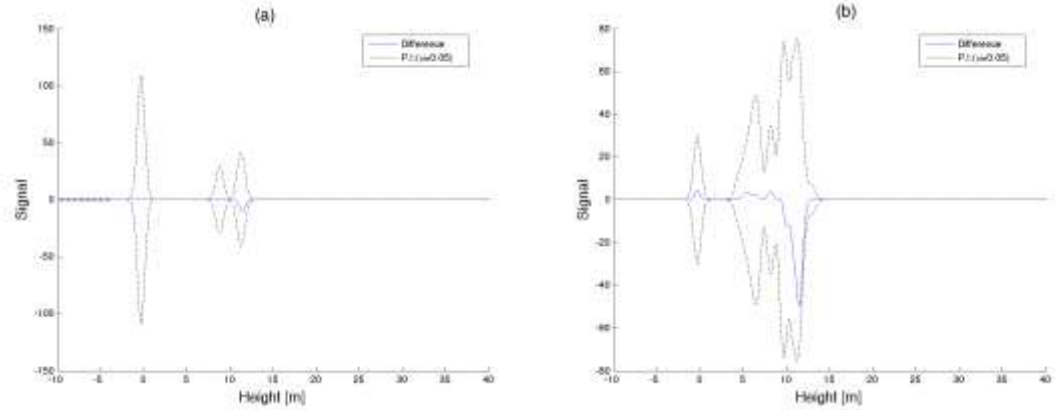


Figure 3: Difference between tb123l (no twigs or leaf stems) and tb123wsl (all) geometries in blue. Prediction interval for the difference with α=0.05 in dashed black. Left is site 1, right is site 2. Both are 4ns outgoing pulse width. Both figures show examples of where the lower geometry waveform falls within the prediction interval.

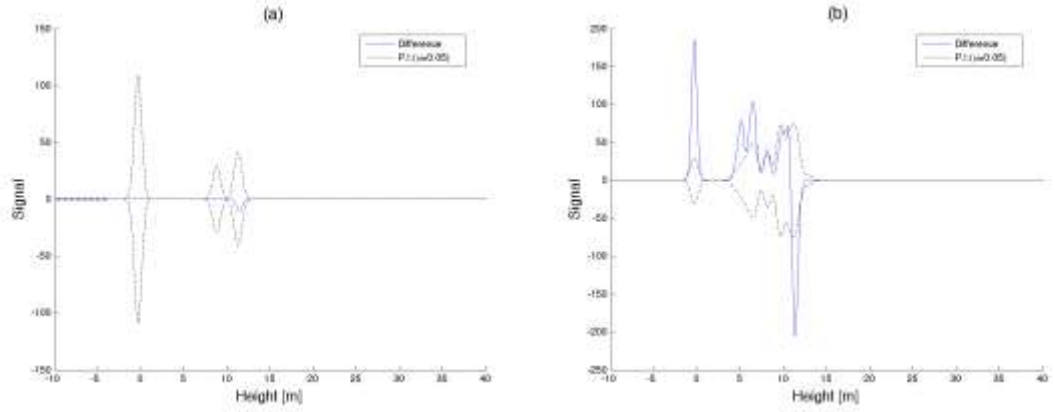


Figure 4: Difference between l (leaf only) and tb123wsl (all) geometries in blue. Prediction interval for the difference with α=0.05 in dashed black. (a) is site 1, (b) is site 2. Both are 4ns outgoing pulse width. Site 1 has the waveform fall inside the prediction interval. Site 2 has part of the waveform fall outside the prediction interval.

Figure 5 shows the relationship between the SAM (eq. 1) values and different levels of geometry. As the geometry complexity was decreased, the difference between the waveforms, as calculated by SAM, increased. This was true for all outgoing pulse widths. The slight decrease in SAM values for site 0 between the tb1l and tbl geometries was attributed to an artifact of SAM (small changes at high vector component values change SAM more than small changes at low values). The SAM values were all small (<2.5°) for the tb123wl and tb123l geometries. The SAM values decreased as the outgoing pulse width increased.

4. Discussion

The difference between the tb123wl and tb123l geometries and the associated validation waveform was not statistically different from zero for all of the sites and all of the outgoing pulse widths (see Figure 3 for an example). The maximum spectral angle between the validation waveform and one of these reduced geometry waveforms was approximately 2.5°. The remaining geometry sets (l, tbl, tb1l, and tb12l) were statistically different from zero for at least one of the sites (e.g., Figure 4).

From visual inspection of Figure 2, any of the waveforms from reduced geometries have a similar shape to the validation geometry. The removal of a component from a specific geometry will affect the resulting signal to a larger degree, as the complexity of the geometry for that given waveform increases. This is shown most clearly in site 2 (Figure 2b). The SAM values for all of the geometries are closer to being the same (0°) than being different (90°). This is due to the wider pulse width having a smoothing effect of the different geometries.

From visual analysis of the waveforms at different geometry levels (Figure 2), the removal of components causes a shift in backscatter energy towards the ground. This is due to more photons being able to propagate further from where the previous geometry was. There was one site (0, not shown) where there was no ground component to the signal until only the leaf geometry was used. This was due to a bough blocking the photons from reaching the ground.

The implications of this study are:

1. It will be nearly impossible to detect objects on the size order of leaf stems or twigs using an aerial LiDAR system that is similar to the simulated system.
2. The tb123l subset will produce consistent waveforms to the validation waveforms in simulations similar to the one listed. This will save disk space and increase run times,

without sacrificing signal fidelity.

3. Decreasing the outgoing pulse width will increase the potential of smaller objects to impact the waveform signal.

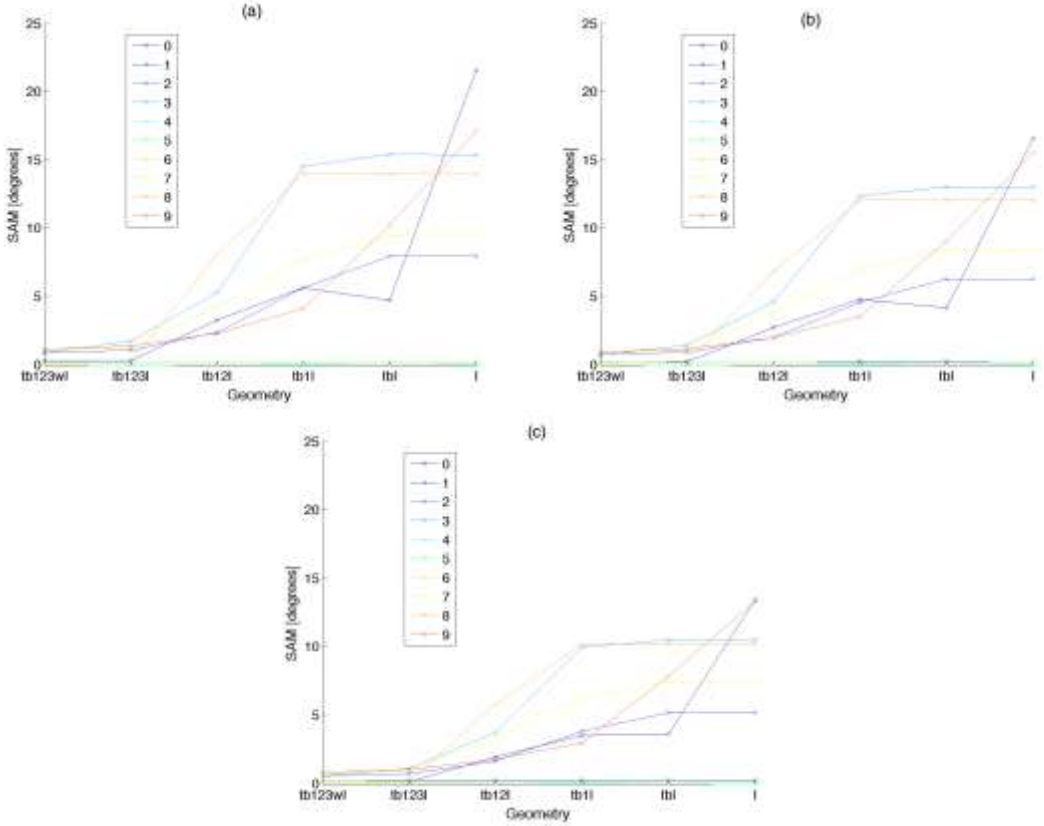


Figure 5: SAM values for the different sites between the mean of the validation waveform and the waveform at a given geometry. (a) is 4ns, (b) is 8ns, and (c) is 16ns outgoing pulse width. The connected lines are for visualization purposes and do not represent intermediate geometries. The SAM values decrease with increasing pulse width and with increasing geometry.

The estimate of the standard deviation of the validation waveform only took into account the random distribution of photons. It did not take into account real world sources of variations such as aligning the sensor target geometries between flights, electronic noise, and differences in the canopy geometry due to the wind. It should be noted that we do not claim that a detection algorithm can be produced to detect any one of the components individually, only that some combination of all of them influence the waveforms.

This survey was only performed for nadir viewing conditions. It is not expected that small deviations from nadir will affect the results significantly. When the viewing angle approaches 90° off nadir, it is expected that the effect the woody parts of a tree have on the resultant waveform will increase, although this has to be verified.

5. Conclusions

The changes in a waveform signal, caused by removal of twigs and leaf stems from a full geometry set, are not statistically significant at a 95% confidence level. Trunks, boughs, branch-level-1, branch-level-2, branch- level-3, and leaves can have a statistically significant impact on a waveform LiDAR signal. A narrower outgoing pulse width can allow for smaller changes in geometry to be observed in a waveform signal, while small changes in the upper

canopy may have an impact on the rest of the waveform signal.

This study shows that we cannot detect structures with similar properties to twigs and leaf stems using a system similar to the one modeled, since they do not cause a statistically significant change to the waveform signal. Future systems with some combination of a smaller footprint, other wavelength(s), faster digitization, and smaller outgoing pulse widths may be able to detect these changes. Additional work will include increasing the number of sites to get better statistics and simulating off-nadir waveforms. Additional geometry sets such as leave-one-out and using only one type of geometry may be used in future simulations to better understand the impact of the individual components. Additionally, putting targets of varying size, shape, position, and spectral properties under the canopy will help to understand the impact of a single component as opposed to the group of components from this study.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1127728. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

C. Mallet and F. Bretar. 2009. Full-waveform topographic lidar: State-of-the- art. *ISPRS J. Photogrammetry and Remote Sensing*, vol. 64, no. 1, 1–16.

Sun, G. and Ranson, K. 2000. Modeling lidar returns from forest canopies. *IEEE Transactions on Geoscience and Remote Sensing*. vol. 38 no. 6, 2617-2626

Ni-Meister, W., and Jupp, D., and Dubayah, R. 2001. Modeling lidar waveforms in heterogeneous and discrete canopies. *IEEE Transactions on Geoscience and Remote Sensing*. vol. 39. no. 9, 1943-1958

Munger W. and Wofsy S. 1999. EMS - Biomass Inventories. *Harvard Forest Data Archive*: HF069.

Bosanac, B. and Zanchi, P. 2011. OnyxTREE BROADLEAF 7.0 http://www.onyxtree.com/

Schott, J., Brown, S., Raqueño, R., Gross, H. and Robinson, G. 1999. An advanced synthetic image generation model and its application to multi/hyperspectral algorithm development. *Canadian Journal of Remote Sensing*, vol. 25, no. 2, 99-111.

Brown, S., Blevins, D, and Schott, J. 2005. Time-gated topographic LIDAR scene simulation. *Proc. SPIE, 5791*. 342-353

Richards, J. and Jia, X. 2006. Interpretation of Hyperspectral Image Data in Remote Sensing Digital Image Analysis: An Introduction 4th ed. Springer. Berlin, Germany: 368-388.

Schott, J., 2007. Detector-Sensor Performance Calculations in Remote Sensing: The Image Chain Approach 2nd ed. *Oxford University Press*. New York, New York: 180-190.

Identification of forest species using fusion of multi-source remote sensing data

Jili Li & Baoxin Hu

Department of Earth and Space Science and Engineering, York University, 4700 Keele Str., Toronto, Ontario, Canada. Email: ljili@yorku.ca; baoxin@yorku.ca.

Paper Number: SL2012-158

Abstract

Tree species identification has been a key research topic to date, while the performance of identification is limited. With the aid of remote sensing technology and multiple space-borne and airborne data, species identification could be improved by using the information fusion technique. This paper proposed an information fusion approach for forest species identification using multi-scale segmentation methods and Dempster-Shafer theory. The principal objective of this paper is to develop a framework of species classification using information derived from combined raster images and LiDAR data based on the Dempster-Shafer theory. We tested the approach for the dominant species in north Ontario forests at tree-group scale and our result demonstrated the usefulness of LiDAR data and Dempster-Shafer theory in the species classification application. An overall accuracy of around 80% was obtained using independent training and testing samples.

1. Introduction

Accurate identification of forest species is critical for forest inventory and sustainable forest management. For example, the estimation of biomass, carbon content, and species diversity requires precise species information. Because of the complexity of forest structures, species identification using optical images has not been completely developed. The emergence of new generation remote sensing systems, especially the very high-resolution digital camera sensors and small-footprint light detection and ranging (LiDAR) systems, can provide more accurate and efficient characterization of forest structures (Wulder *et al*., 2012). Studies on identifying species of forest stands or individual trees have been successfully conducted using either airborne/satellite images (Leckie *et al*. 2003; Mora *et al*. 2010; Franklin *et al*., 2000) or LiDAR data (Holmgren and Persson, 2004; Korpela *et al*., 2010). However, remote sensing data acquired over the same forest area by different sensors are generally redundant but complementary, because different sensors measure different physical properties of forests/trees at different spatial scales. Fusion of complementary data provides complete description of a complex forest scene, while there are also conflicts due to differences in resolution, feature distribution, and object scales, which bring us a substantial difficulty for the information fusion. It is extremely important to develop methods to accurately characterize forest scenes and improve species classification accuracy using multi-source information.

Object-oriented methods have been widely developed in remote sensing and forestry applications (e.g., Chen *et al*., 2012; Blaschke 2010). In multi-sensor data fusion analysis (e.g., Heinzel and Koch, 2012; Waser *et al*., 2011), images from multi-sensors are resampled and interpolated based on the highest resolution such that each object covers exactly the same number of pixels on the resampled image. In their ways, single-scale segmentation on the image which has the highest spatial resolution is sufficient for the complete dataset. However, resampling methods normally alter the original spectral features that can be derived from low resolution images, and the representation of small objects such as individual trees is

meaningless. For instance, one pixel on an Advanced Spaceborne Thermal Emission and Reflection (ASTER) image characterizes the compound reflectivity of several trees. Using the gray value of the pixel as a feature of a single tree would increase the complexity for classification and interpretation. It is generally understood that low resolution remote sensing image can characterize forest canopy structures such as homogeneity at large spatial scale, while high resolution image and LiDAR data can characterize crown structures at individual-tree or tree-group scale. To effectively apply the information, multi-scale segmentation is one of the best ways to handle multiple datasets from different sources.

The Dempster-Shafer theory (DST) was initially proposed by Dempster (1968) and Shafer (1976) as a Bayesian framework allowing fusion of information from independent sources. For an object, the theory does not require direct features derived from different data as other hard-classification, and allows the consideration of uncertainty and vagueness based on knowledge. As a mean of converting feature values to probabilities, DST is an appropriate methodology for the information fusion scenario in our classification theme. Previous studies have demonstrated strong potentials of DST in classification applications (Rottensteiner *et al*., 2005). It is also applied in forest applications such as mapping stand healthy of mature deciduous forests (Mora *et al*., 2012), classifying landscape objects (Cayuela, *et al*., 2006), and mapping regenerating stands (Mora *et al*., 2011), but it is rarely used for individual species classification. Objective of this study is to classify the dominant species in north Ontario forests based on tree-group scale (i.e., 5-20 grouped trees) using information derived from combined raster images and LiDAR data, and to test the usefulness of DST in the forest species classification application.

2. Methods

2.1 Materials

Located in the Great Lakes-St. Lawrence forest region near Sault Ste. Marie, Ontario, Canada, our study area (Figure 1) included four forest sites (S1, S2, S3 and S4) with areas ranging from about 50 to 225 ha. These sites were selected based on road access, various species and stand structural conditions represented. The dominant species at these sites are eastern white pine (*Pinus strobes* L.), jack pine (*Pinus banksiana* Lamb.), trembling aspen (*Populus tremuloides* Michx.), and sugar maple (*Acer saccharum* Marsh.). Occasionally, white birch and black spruce trees can be found mixed with the dominant species. The species in the largest site S1 is consisted of eastern white pine, jack pine, and deciduous trees. The jack pine trees in this site are mostly homogenous plantations that formed with even-aged stand structure. S2 is dominated by white pine (about 90% stocking) and S3 is dominated by sugar maple (about 40% stocking) and co-dominated by trembling aspen (about 35% stocking). The tree species in the three sites S1-S3 are evenly distributed and generally not mixed with each other, while maple, aspen, and a few coniferous trees are seriously mixed in the last site S4. Forest disturbances in the study area are likely due to harvesting, insects, and scientific research. The forests can be characterized by different homogeneous and inhomogeneous stands according to the tree age, species composition, and vertical structure.

Figure 1: Location (left) and QuickBird image (right) of study area, and relative location of four study sites (S1-S4).

Several remote sensing datasets are available for the four study sites including: ASTER, high-resolution QuickBird image, very high-resolution aerial images obtained from digital cameras, and discrete LiDAR point cloud. Detail technical characteristics for these datasets are given in Table 1. In this study, we used the green, red and near infra-red bands of the ASTER imagery. Except the dataset presented in Table 1, another 0.15 m resolution digital images are captured simultaneously with the LiDAR data and they are used together with field data as references to interpret species for classification validation. To keep the original spectral information, none of the images is resampled and filtered. The discrete LiDAR data are classified into ground points and non-ground points and they are used to create a digital terrain model (DTM) and a digital surface model (DSM). A raster format canopy height model (CHM) is calculated as the difference between DSM and DTM. The CHM is smoothed using a low pass filter and post-processed to remove abnormal height pixels.

Table 1: Summary of characteristics of the image and LiDAR data used for species classification

Data	ASTER	QuickBird	Aerial image	LiDAR
Acquisition time	2007/07	2008/06	2008/09	2009/08
Spectral band	G, R, NIR	R, G, B, NI	R, G, B, NI	-
Pixel resolution	15 m	2.4 m	0.4 m	-
Radiometric resolution	8 bit	8 bit	14 bit	-
Point density	-	-	-	about 40 points m-2
Number of echo types	-	-	-	4

Field data over the four study sites were collected during August 2009 and the sites were revisited in August 2011. During the field campaigns, the dominant and co-dominant species and stand structure information of each site were recorded and overlapped with the 0.15 m high-resolution camera image. Besides the species at stand level, several individual trees in each site were also selected representatively as sample trees for each species type and their location, height, and diameter at breath height were measured.

2.2 Dempster-Shafer theory

The DST firstly defines a frame of discernment Θ that contains all the θ_i classes under consideration: $\Theta = [\theta_1, \theta_i, ..., \theta_N]$. A power set 2^Θ is then defined including all the subset of Θ and the empty set ϕ. For example, if N=2, the power set is:

$$2^\Theta = [\theta_1, \theta_2, \theta_1 \bigcup \theta_2, \phi]. \quad (1)$$

DST utilizes mass functions characterizing the confidence of a evidence belong to each focal element (e.g., θ_1). The mass function $m(A)$ are in a form of probabilities such that $0 \le m(A) \le 1, m(\phi) = 0$, and $\sum_{A \in 2^\Theta} m(A) = 1$, where $A \in 2^\Theta$. Let n defines the number of evidences, probability masses $m_i(B_j)$ can be defined for all classes $B_j \in 2^\Theta$. DST allows the combination of these probabilities from multiple evidences to calculate a combined mass for each class $A \in 2^\Theta$:

$$m(A) = \frac{\sum_{B_1 \bigcap B_i ... \bigcap B_n = A} \prod_{1 \le i \le n} m_i(B_j)}{1 - \sum_{B_1 \bigcap B_i ... \bigcap B_n = \phi} \prod_{1 \le i \le n} m_i(B_j)} \quad . \quad (2)$$

Two decision rules supporting credibility and plausibility can be defined for all $A \in 2^\Theta$:

$$Cr(A) = \sum_{B_i \subseteq A} m(B_i), \quad (3)$$

$$Pl(A) = \sum_{B_i \in 2^\Theta, A \bigcap B_i \ne \phi} m(B_i). \quad (4)$$

For each decision rule, the hypothesis which maximizing the decision statistics is adopted as the most credible or plausible decision (More *et al*., 2011). In this study, we used the maximum credibility.

2.2 Pre-process and segmentation

Generally, object-oriented fusion of remote sensing data can be performed at one of the three stages: data processing stage, segmentation stage, and classification stage. In this study, we put the fusion method in the last stage (i.e., classification). Figure 2 shows a workflow of the methodology implemented. For each study site, the 15 m resolution ASTER image (G, R, and NIR bands), the 2.4 m resolution QuickBird image (RGB and NI bands), the 0.4 m resolution aerial image (RGB and NI bands), and the 1 m resolution CHM image were first geometrically co-registered together to ensure that the error due to mismatching was within 1 pixel. Next, the CHM was resampled to 2.4 m to geographically overlap with the QuickBird and aerial images.

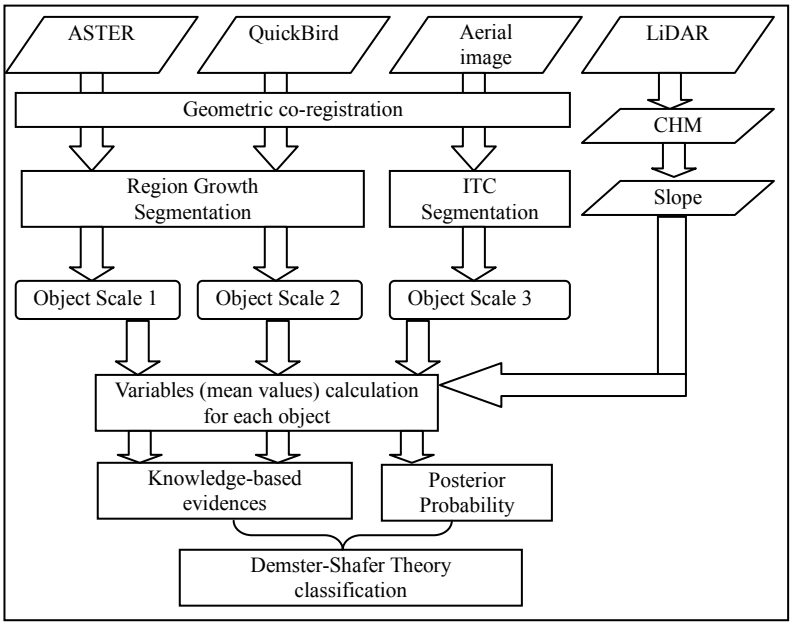


Figure 2: Workflow of forest species classification based on Dempster-Shafer theory.

Two segmentation algorithms were applied to delineate images into meaningful objects at multi-scale. The first algorithm used is Region Growing implemented by the SPRING 5.2 software (SPRING - DPI/INPE, 2012). We used the segmentation module in SPRING 5.2 to process the ASTER and QuickBird images. For each of the two datasets, three spectral bands G, R, and NIR for ASTER, and R, G, B for QuickBird images were used as input raster layers. The two parameters, similarity and area (pixels) in the region growing algorithm were determined empirically based on trial-and-error. The best parameter combination for ASTER and QuickBird images are: [Similarity=2, Area=10] and [Similarity=1, Area=100], respectively. The second algorithm used is a multi-scale individual tree crown (ITC) segmentation method derived in our previous study (Jing et al., 2012). Although this algorithm was developed for delineating individual tree crowns, boundary of tree-groups can be still obtained by changing the scale parameters of crowns to larger values (Jing et al., 2012). The scale parameter used to segment aerial image based on the ITC algorithm (Jing *et al*., 2012) was [23, 33, 43], where 33 represents the dominant segment size (diameter in pixels).

In this way, forests were segmented at a coarse level based on ASTER image, at a middle level based on Quickbird image, and at a fine level based on aerial image. The segments in ASTER image represent large areas (canopy) with relative homogenous forest; the segments in QuickBird image represent middle area (stand) homogenous forests; and the segments in aerial image represent small area homogenous tree-groups.

Besides the original RGB and NI image layers, following raster layers were also created for all data sources using ENVI software: (1) NDVI for QuickBird and arieal data, (2) gray level co-occurrence matrix (GLCM) based homogeneity feature derived based on the green band of each data source using ENVI software, and (3) slope using CHM.

2.3 Selection of species categories

In this study, classification of forest species was tested at tree-group scale. Objects were defined as the obtained segments on aerial image using ITC algorithm and large scale parameters. An individual objects normally covered a group of 5-20 individual trees depending on the spatial

distribution and size of trees. The species classes to be evaluated in this study are: white pine (Pw), jack pine plantation (PjP), sugar maple (Ms), trembling aspen (Pt), and mixedwood (Wm).

2.4 Dempster-Shafer classification

The DST is a generalization of Bayesian theory of probability that handles the combination of multiple independent evidences derived from various sources. Using DST, the degree of belief for different classes or grouped classes can be obtained. For the details of the DST, especially in the forest applications, see Mora *et al*. (2011). Here, we mainly present the construction of probability masses.

Segments obtained based on aerial image have unique object identities (single integers). Given an object, we calculated the mean values of the RGB and NI bands from the aerial image. A Navie Bayes classification algorithm (Mitchell, 1997) was then used to calculate the posterior probabilities for the given object belonging to each class. The Navie Bayes classifier assumes that the presence of a particular feature of a class is unrelated to the presence of any other features, and all the features independently contribute to the probability distribution for a given class. The posterior probabilities were then used as one of the mass functions for the aerial image.

For the ASTER, QuickBird, and CHM data, mass functions were assigned manually, and were based on expert knowledge. They are used to support singleton or compound hypotheses (Table 2). The evidences were designed based on: (1) canopy spectral feature at red band, (2) canopy homogeneity, (3) stand homogeneity, (4) stand forest cover - NDVI, and (5) stand surface slope. They are summarized in the following descriptions.

1. Canopy spectral feature at red band refers to the mean value of all pixels belong to the object. It is obvious based on ASTER data that coniferous trees generally has lower green pixel value than deciduous trees, and a threshold of DN=45 was used to easily separate coniferous and deciduous trees in our study area. This evidence supports class group: [Pw, PjP]. For each object, we assigned the probability mass $P_1 = \begin{cases} 0.7, x \le 45 \\ 0.3, x > 45 \end{cases}$, where x is the mean value of pixels of the red band.
2. Canopy homogeneity refers to the mean value of GLCM homogeneity features belong to the object. It is calculated based on ASTER data. This evidence supports the jack pine plantation and deciduous trees [PjP, Ms, Pt]. The object with evidence value close to 1 represents homogenous canopy. We assigned the probability mass $P_2 = \begin{cases} 0.7, x > 0.7 \\ 0.3, x \le 0.7 \end{cases}$, where x is the mean value of GLCM homogeneity features of the object.
3. Stand homogeneity has the similar meaning as canopy homogeneity, except that the stand homogeneity is derived from QuickBird data. Similarly, we assigned the probability mass $P_3 = \begin{cases} 0.7, x > 0.7 \\ 0.3, x \le 0.7 \end{cases}$, where x is the mean value of GLCM homogeneity features of the object.
4. NDVI is often used for separating vegetation and non-vegetation objects, while in this study we found that it is useful for classifying coniferous and deciduous trees. Coniferous trees exhibit lower NDVI values than deciduous trees. This evidence

supports class group [Pw, PjP]. We assigned the probability mass $P_4 = \begin{cases} 0.7, x \le 0.7 \\ 0.3, x > 0.7 \end{cases}$, where x is the mean NDVI value of the objects derived from QuickBird data.

5. Stand surface slope refers to the slope variation of crown tops in the stand. White pine and mixedwood stands have larger slope than the other species types. Due to the even-aged condition, jack pine plantation tends to have similar variation as other deciduous trees. This evidence supports the classes group: [Pw, Wm]. We assigned the probability mass $P_5 = \begin{cases} 0.7, x > 35 \\ 0.3, x \le 35 \end{cases}$, where x is the mean slope value of objects from the slope raster layer derived from CHM.

Many decision rules have been proposed for choosing the best hypothesis (classes) after the combination of evidences. In this study, we applied the most common decision rules defined by Shafer (1976): the maximum credibility and maximum plausibility.

Table 2: Evidences in support of different class groups used in the species classification using DST. Pw: white pine; PjP: jack pine plantation; Ms: sugar maple; Pt: trembling aspen; Wm: mixedwood.

Data source	Evidence	Supported class group	Mass function type	Probability range
ASTER: Red	Canopy spectral feature at green band	[Pw, PjP]	Fixed probability	0.3/0.7
ASTER: Texture	Canopy homogeneity	[PjP, Ms, Pt]	Fixed probability	0.3/0.7
Aerial image: RGB+NI	Native Bayesian posterior probability	[Pw], [PjP], [Ms], [Pt], [Wm]	Posterior probability	0.0-1.0
LiDAR: Slope	Stand surface slope	[Pw, Wm]	Fixed probability	0.3/0.7
QuickBird: NDVI	Stand forest Cover	[Pw, PjP]	Fixed probability	0.3/0.7
QuickBird: Texture	Stand homogeneity	[PjP, Ms, Pt]	Fixed probability	0.3/0.7

3. Results and discussion

The amount of segments derived for the four study sites S1-S4 were: 2266, 265, 579, and 239, respectively. As an example, the segmentation result of site S2 is shown in Figure 3. Among those segments (objects), we selected 322 samples covering the five species categories as training and testing dataset. For each species, we selected the samples to ensure that any of the training samples and any of the testing samples are not located within a unique object in the large scale segmentation map derived from ASTER. As a result, 161 samples were selected for training and another 161 samples for testing. The overall accuracy of classification based on Navie Bayesian using only the RGB and NI bands of aerial image was 72.1%. Using the DST, the overall accuracy was increased to 80.8%. The number of correctly identify tree-groups were increased mainly for white pine and mixedwood (Table 3 and Table 4), and the user's accuracies for individual species were also improved. The jack pine plantation was always identified with 100% accuracy, which is not surprising for us because their spectral signatures and canopy shape are extremely homogeneous and unique. Table 3 and Table 4 also indicate that sugar maple is the most difficult species to be identified, likely due to the varied range of ages and non-homogenous crowns. The sugar maple trees in our study area sometimes are growing mixed with oak trees, resulting in various spectral signatures. We also compared the classification accuracy using all datasets to the accuracy without LiDAR data. Our results indicate that DS classification using combined LiDAR and optical data can be improved with about 3% overall accuracy, especially for the separation of white pine and other species.

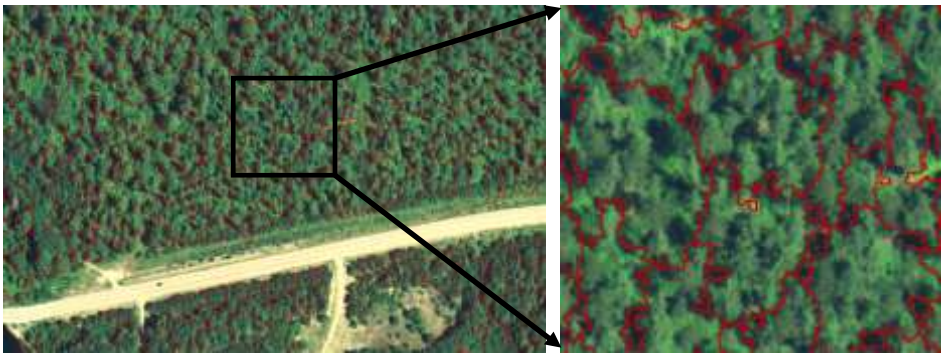


Figure 3: Segmentation result at tree-group scale on study site 2 (left) and a subset demonstration (right).

Table 3: Error matrix of Navie Bayesian classification at tree-group scale using only aerial image. Pw: white pine; PjP: jack pine plantation; Ms: sugar maple; Pt: trembling aspen; Wm: mixedwood.

	Reference					
	Pw	PjP	Ms	Pt	Wm	User's accuracy
Pw	54	0	2	0	5	88.5%
PjP	0	32	1	0	0	100.0%
Ms	5	0	12	2	14	36.4%
Pt	0	0	2	6	0	75.0%
Wm	1	0	9	4	12	46.1%
Total	60	32	26	12	31	
Producer's accuracy	90%	100%	46.2%	50%	38.7%	
Overall accuracy	72.1%					

4. Conclusions

In this paper we proposed a Demster-Shafer theory based classification method using multi-source remote sensing data to classify dominant species of tree-groups. Throughout the proposed method, spectral information from different sensors was naturally combined together in a form of mass probability. Multi-scale segmentation algorithms provide the feasibility of geographic link of an object among different spatial resolutions without resample original datasets. About 80% overall accuracy of species classification can be obtained in our study area using the proposed method. Combining imagery and LiDAR data is also proved effective to increase species classification accuracy. However, the methodology still needs to be refined and tested on large forest area before it can be applied to large forests operationally. In the future, more efforts should be made on the assignment of mass functions as well as a set of sensitivity analysis.

Acknowledgements

The authors would like to thank Natural Sciences and Engineering Research Council of Canada (NSERC) for supporting the research, thank Geodigital Inc. for providing the LiDAR data, and thank Ontario Ministry of Nature Resource (OMNR) for providing the high-spatial resolution aerial images.

Table 4: Error matrix of Damster-Shafer classification at tree-group scale using combined image and LiDAR datasets. Pw: white pine; PjP: jack pine plantation; Ms: sugar maple; Pt: trembling aspen; Wm: mixedwood.

	Reference					
	Pw	PjP	Ms	Pt	Wm	User's accuracy
Pw	59	0	1	0	4	92.1%
PjP	0	32	0	0	0	100%
Ms	1	0	16	2	11	53.3%
Pt	0	0	1	7	0	87.5%
Wm	0	0	8	3	16	59.3%
Total	60	32	26	12	31	
Producer's accuracy	98%	100%	61.5%	58.3%	51.6%	
Overall accuracy	80.8%					

References

Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 2-16.

Chen, G., Hay, G. J., and St-Onge, B., A GEOBIA framework to estimate forest parameters from lidar transects, Quickbird imagery and machine learning: A case study in Quebec, Canada. *International Journal of Applied Earth Observation and Geoinformation*, 15, 28-37.

Cayuela, L., Golicher, J. D., Rey, J., S., and Benayas, J. M. R., 2006. Classification of a complex landscape using Dempster-Shafer theory of evidence. *International Journal of Remote Sensing*, 27(10), 1951-1971.

Dempster, A. P., 1968. A generalization of Bayesian inference. *Journal of the Royal Statistical Society*, Series B, 30, 205-247.

Franklin, S. E., Hall, R. J., Moskal, L. M., Maudie, A. J., & Lavigne, M. B. (2000). Incorporating texture into classification of forest species composition from airborne multispectral images. *International Journal of Remote Sensing*, 21, 61-79.

Holmgren, J., and Persson, Å., 2004. Identifiying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, 90, 415-423.

Heinzel, J., and Koch, B., 2012. Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *International Journal of Applied Earth Observation and Geoinformation*, 18, 101-110.

Jing, L., Hu, B., Noland, T. and Li, J., 2012. An individual tree crown delineation method based on multi-scale segmentation of imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 70, 88-98.

Korpela,I., Ørka, H. O., Maltamo, M., Tokola, T., and Hyyppä, J. (2010). Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, 44, 319-339.

Leckie, D. G., Gougeon, F. A., Walsworth, N., and Paradine, D., 2003. Stand delineation and composition estimation using semi-automated individual tree crown analysis. *Remote Sensing of Environment*, 85, 355-369.

Mitchell, T. M., 1997. Machine Learning, McGraw-Hill Science.

Mora, B., Wulder, M. A., and White, J. C., 2010. Idendifying leading species using tree crown metrics derived from very high spatial resolution imagery in a boreal forest environment. *Canadian Journal of Remote Senisng*, 36(4), 332-344.

Mora, B., Fournier, R., and Foucher, S., 2011. Application of evidential reasoning to improve the mapping of regenerating forest stands. *International Journal of Applied Earth Observation and Geoinformation*, 13, 458-467.

Mora, B., Fournier, R., and Foucher S., 2012. Mapping the health of mature deciduous forest stands by fusing multisource geospatial data with Dempster's combination rule. *International Journal of Remote Sensing*, 33(4), 1139-1163.

Rottensteiner F., Trinder, J., Clode, S., and Kubik, K., 2005. Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. *Information Fusion*, 6, 283-300.

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton, NJ, Princeton University Press.

SPRING 5.2 - DPI/INPE, 2012, http://www.dpi.inpe.br/spring, last accessed on August 1, 2012.

Waser, L. T., Ginzler, C., Kuechler, M., Baltsavias, E., and Hurni, L., 2011. Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. *Remote Sensing of Environment*, 115, 76-85.

Wulder M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Hilker, T., Bater, C. W., and Gobakken, T., 2012. LiDAR sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, 121, 196-209.

Incorporating LiDAR-derived tree height data into an Alberta Vegetation Inventory (AVI) operational softcopy workflow

Michael Chubey[1], Geoff Lawless[1] Maurice Lapierre[1] & Kevin Stehle[1]

[1]Silvatech Consulting Ltd. m.chubey@silvatech.ca

Paper Number: SL2012-161

Abstract

The utility of LiDAR data for generating a wide range of forest inventory parameters has been demonstrated time and again by many researchers from around the globe over the course of the last decade. Yet adoption of LiDAR technology for use in large scale operational forest inventory projects by non-researchers has been slow, possibly partially due to the lack of practical tools and methods designed for integrating LiDAR data into the operational forest inventory process. The conventional approach for compiling forest inventory information is though the process of photo interpretation using techniques that have changed little over the last half century. In an attempt to reduce the subjectivity inherent in the photo interpretation process with respect to tree height related attributes, an approach was developed for incorporating automatically generated LiDAR-derived tree height profile data into the operational softcopy photo interpretation workflow of a large forest inventory project. We feel that the objective nature of the approach described herein translated into greater consistency in tree height related attributes thus increasing the integrity of the final inventory product.

1. Introduction

Forest vegetation inventories in most jurisdictions in Canada are compiled by certified photo interpreters in a photogrammetric or softcopy environment in conjunction with limited ground sampled calibration data (Pinto *et al*., 2007; Woods *et al*., 2011). While improvements have been made to the overall framework, the underlying approach is the same as that developed in the 1950's and relies heavily on the skill and experience of individual interpreters (Leckie and Gillis, 1995; King, 2000). The recent shift to softcopy photo interpretation from hardcopy delineation on photographs has led to incremental improvements to delivery times, but has not resulted in substantial improvements to the information quality or reliability of vegetation inventories (Pinto *et al*., 2007; Magnussen and Russo, 2012).

Attributes related to tree height and vertical stand structure are key components of any forest resource inventory. Many studies over the past decade have documented the utility of LiDAR for generating height related forest inventory attributes (Lim *et al*., 2003; Naesset, 2004; Zhao *et al*., 2011). To date, however, examples of LiDAR being used in the production of forest inventories in an operational setting, particularly in Canada, are scarce (Hilker *et al*., 2008; Woods *et al*., 2011; Treitz *et al*., 2012). The slow adoption of LiDAR to operational inventory applications has largely been attributed to cost; however, another reason may relate to the fact that tools and methods for integrating LiDAR data into real-world operational inventory workflows are not readily available. The purpose of this paper is to describe an approach for incorporating LiDAR-derived tree height data into an operational softcopy workflow and discuss how it was applied in a recent large scale Alberta Vegetation Inventory (AVI) project.

2. Methods

The project area was located in the Boreal Forest Natural Region of Alberta, Canada, approximately 350 km northwest of Edmonton and covered an area of approximately 400,000 ha (Figure 1). Landcover was mainly forested with species comprised primarily of trembing aspen (*Populus tremuloides*), black spruce (*Picea mariana*), white spruce (*Picea glauca*), and jack pine (*Pinus banksiana*), with minor amounts of tamarack (*Larix laricina*), balsam poplar (*Populus balsamifera*), paper birch (*Betula papyrifera*), and balsam fir (*Abies balsamea*).

Figure 1: Project area in the Boreal Forest Natural Region of Alberta, Canada.

Original LiDAR point cloud data covering the entire project area with an approximate post spacing of 1 m were processed to produce a point-based canopy height model (CHM), representing height above ground at the original LiDAR point locations. With an original post spacing of ~1 m, it was understood that most of the LiDAR points in the data set represented elevations somewhere between top-of-canopy and the ground. In an attempt to identify points that corresponded as closely as possible to actual tree top measurements (as opposed to mid-canopy hits), the CHM was further processed using a point filtering technique which thinned the CHM data to retain only the highest elevation points at the centre of a moving 15 m diameter neighbourhood (Figure 2). The horizontal distance filtering threshold (15 m in this case) was chosen after comparing various point filtering results with field measured tree heights from 72 plots distributed across the project area. All LiDAR point cloud processing was performed using the Terrascan software package (www.Terasolid.com).

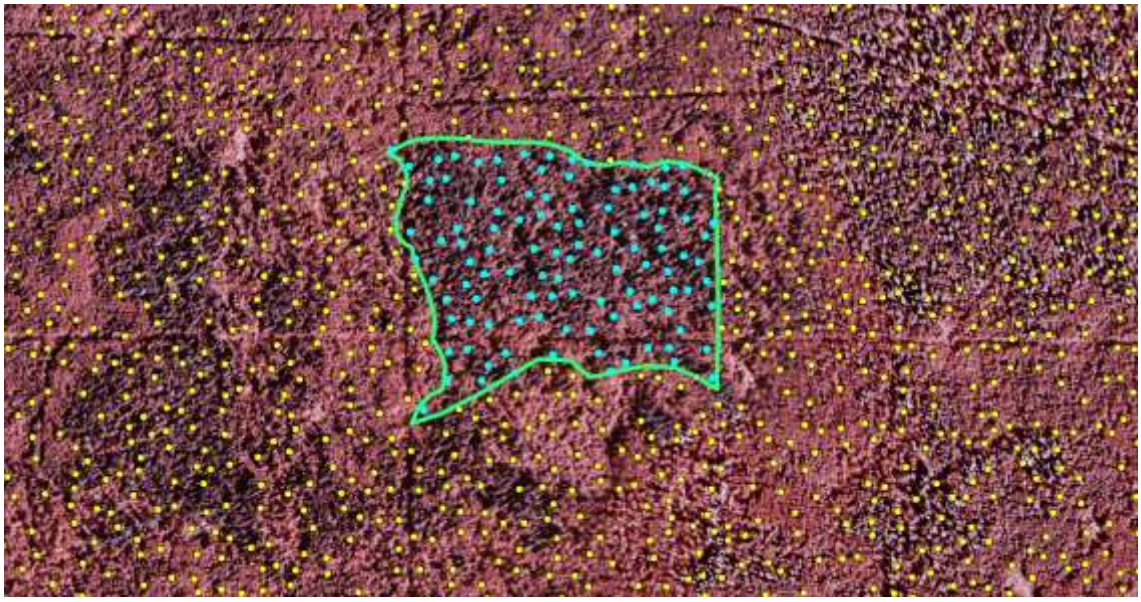


Figure 2: Filtered CHM points representing maximum heights every ~15 m overlaid on colour infrared (CIR) digital imagery. Points falling within a forest polygon are selected (highlighted in blue) and subsequently used to generate a height frequency distribution for the selected polygon (see Figure 3).

A custom statistical analysis software tool was developed and implemented within the softcopy environment which enabled interpreters to utilize the processed LiDAR data "on the fly" by automatically generating height frequency distributions and summary statistics for selected forest polygons. The resulting height frequency histograms were used by the interpreters to quickly assess LiDAR tree height related characteristics for any polygon (Figure 3).

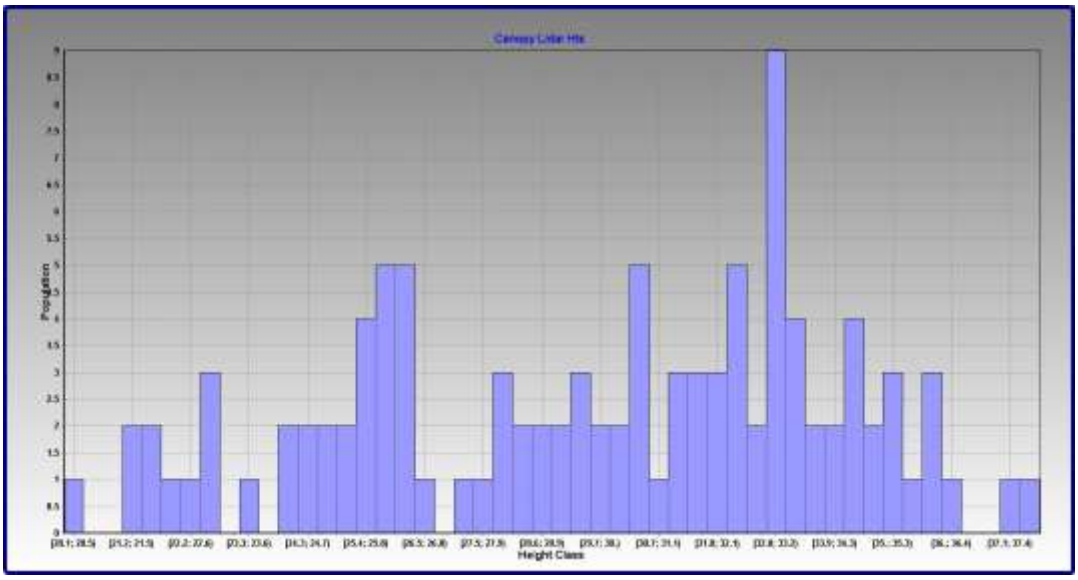


Figure 3: Frequency distribution histogram summarizing LiDAR-derived canopy height data for the forest inventory polygon highlighted in Figure 2.

3. Results and Discussion

Accuracy of the LiDAR-derived heights was assessed by comparing field measured heights with mean LiDAR-derived heights for polygons in which the field plots were located. The relationship between LiDAR-derived polygon heights and corresponding field plot heights is shown in Figure 4 ($R^2 = 0.77$, $p < 0.05$, n=72). A *t*-test found no significant difference between the two data sets ($p > 95\%$). The strong relationship between field measured and LiDAR-derived heights observed in this analysis is consistent with findings from several recent studies (Means, *et al*., 2000; Coops *et al*., 2007; Woods *et al*., 2011).

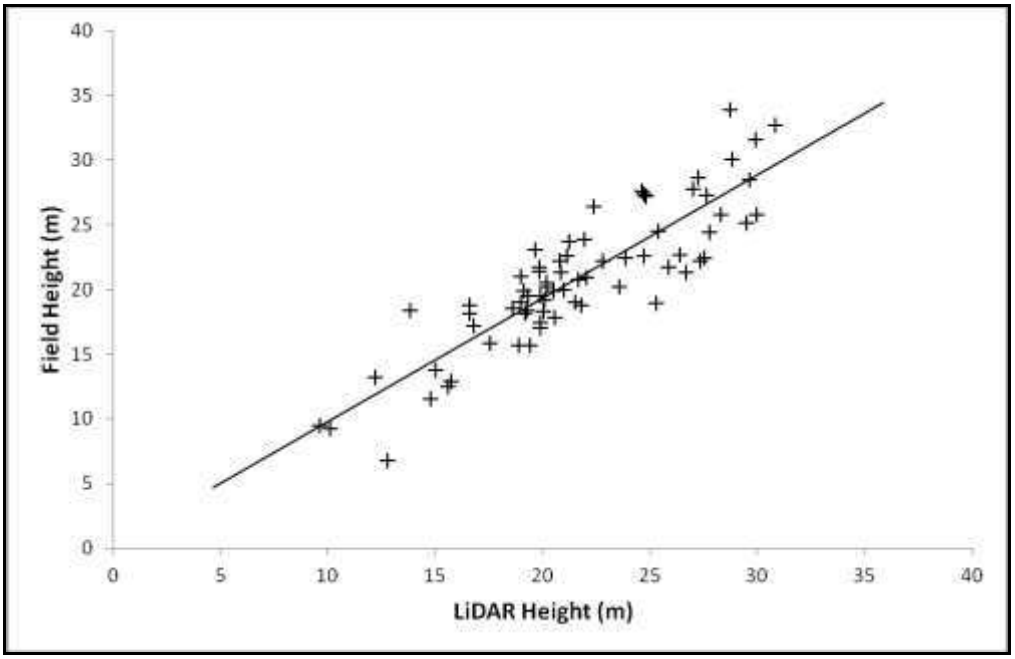


Figure 4. Relationship between LiDAR-derived mean polygon heights and corresponding field measured heights ($R^2 = 0.77$, $p < 0.05$, n=72).

The conventional approach for deriving stand level tree heights in a forest inventory is manual measurement of a small number of trees, thought to be representative of the stand, by a photo interpreter in a photogrammetric or softcopy environment. With this approach, potential subjective bias exists in several areas including the number of trees selected for measurement, the choice of trees to be measured, and the skill with which the trees are measured. By contrast, the LiDAR-based approach applied in this project offers an objective alternative that significantly reduces interpreter bias by almost completely eliminating the need for selecting and measuring trees in softcopy. Once accustomed to evaluating the height profile data, interpreter stand height calls were found to be more consistent compared to what would be expected from the conventional approach.

In addition to stand height, the frequency distribution histograms generated “on the fly” offered interpreters a quick, easy, and accurate method of assessing stand structure and determining whether a stand is single- or multi-story. For example, interpretation of the histogram in Figure 3 reveals a two story stand, characterized by a “double hump” distribution pattern. Softcopy analysis of colour infrared (CIR) imagery revealed a white spruce leading stand with a mature mixed spruce/trembling aspen overstory (90% white spruce) and a younger spruce/aspen understory layer (80% white spruce). In this example, the overstory white spruce canopy layer ranges in height from 30-37 m with a dominant height of approximately 33 m and the sparse

aspen component measuring approximately 27 m. The understory white spruce layer is approximately 25-26 m while the understory aspen layer varies between 21-22 m.

A height frequency distribution for a single story trembling aspen stand is shown in Figure 5. The histogram for this stand exhibits a classic bell shaped distribution pattern characteristic of a single story forest canopy and reveals a dominant stand height of 21 m.

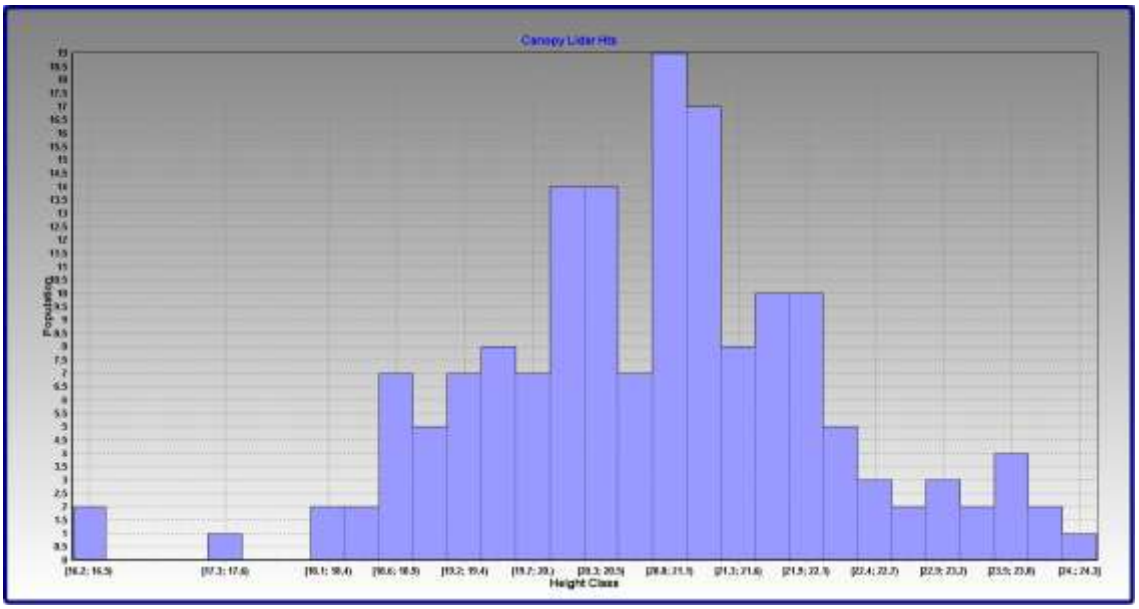


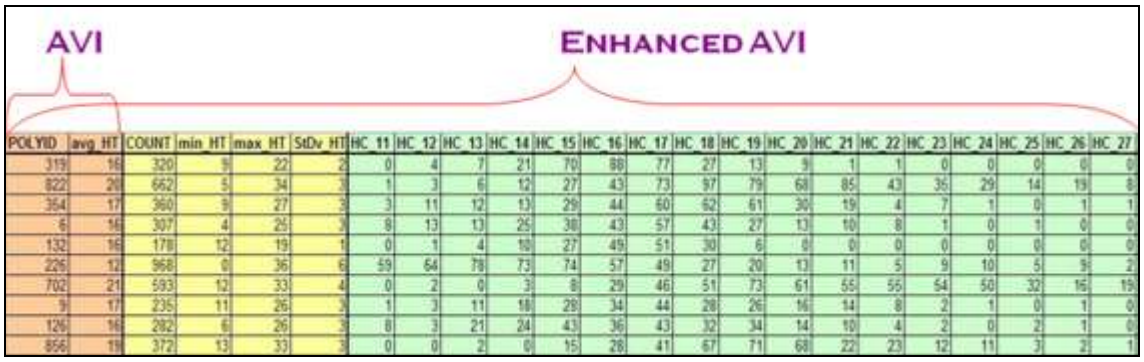
Figure 5: Height frequency distribution for a single story trembling aspen stand.

One limitation of the LiDAR-based histogram height interpretation approach is that it did not work well in wetland complex stands. In these stands, black spruce and larch/tamarack species often grow in very close proximity to one another in dual species canopy layers but with the larch/tamarack trees significantly taller than black spruce trees. Due to the way in which the LiDAR data were processed for this project (i.e. filtering to retain only maximum heights every ~15 m), black spruce trees were not well represented in the LiDAR-derived height frequency distributions. In these cases, interpreters needed to calibrate their height determinations for black spruce trees based on field data and their knowledge of black spruce and larch/tamarack relationships.

The full effect of lidar integration for this project is difficult to quantify since several aspects of the inventory process were affected, sometimes in subtle ways. However, qualitative feedback from interpreters indicated that, in general, they liked the approach and felt that it helped improve the overall integrity of the final product. Several interpreters reported feeling more confident in their height calls which translated into less speculation or "second guessing" with respect to attribution of several other variables including stand age, structure, species mix, and timber productivity rating (TPR). The final inventory product was found to meet or exceed Government of Alberta specifications based on the results of a standard provincial AVI audit.

A by-product of the LiDAR analysis performed in this project was tabular summaries of height frequency distribution data by polygon. This information was appended to the attribute tables of the project deliverable AVI polygons as enhanced tree height attributes (Figure 6). While traditionally, the main application of the forest inventory in Canada has been timber management, it is now accepted that forest inventory information is important for a broad range

of social, environmental, and economic forest management goals (Wulder *et al*., 2008; Special Committee on Timber Supply, 2012). Enhanced information such as the detailed depiction of stand vertical structure shown in Figure 6 could potentially be leveraged for non-traditional purposes such as assessing habitat suitability for particular wildlife species.



POLYID	avg_HT	COUNT	min_HT	max_HT	StDv_HT	HC_11	HC_12	HC_13	HC_14	HC_15	HC_16	HC_17	HC_18	HC_19	HC_20	HC_21	HC_22	HC_23	HC_24	HC_25	HC_26	HC_27
319	16	320	9	22	2	0	4	7	21	70	88	77	27	13	9	1	1	0	0	0	0	0
822	28	662	5	34	3	1	3	6	12	27	43	73	97	79	68	85	43	36	29	14	19	8
354	17	360	9	27	3	3	11	12	13	29	44	60	62	61	30	19	4	7	1	0	1	1
6	16	307	4	26	3	9	13	13	25	38	43	57	43	27	13	10	8	1	0	1	0	0
132	16	178	12	19	1	0	1	4	10	27	45	51	30	6	6	0	0	0	0	0	0	0
226	12	968	0	36	6	59	64	78	73	74	57	49	27	20	13	11	5	9	10	5	9	2
702	21	593	12	33	4	0	2	0	3	8	29	46	51	73	61	55	55	64	50	32	16	19
9	17	235	11	26	3	1	3	11	18	28	34	44	28	26	16	14	8	2	1	0	1	0
126	16	282	6	26	3	8	3	21	24	43	36	43	32	34	14	10	4	2	0	2	1	0
856	19	372	13	33	3	0	0	2	0	15	28	41	67	71	68	22	23	12	11	3	2	1

Figure 6: Enhanced height related information included in the final polygon attribute table.

4. Conclusion

This paper has presented an example of how LiDAR data can be implemented within an operational softcopy workflow as an alternative to conventional photo interpretation-based methods for compiling forest inventory height attributes. The approach of using LiDAR-based, statistically-derived tree height profiles as an aid to assessing height related attributes was found to reduce subjectivity in the interpretation process thus increasing the integrity of the final inventory product. Although much research has proven the suitability of LiDAR for deriving a host of forest parameters, more operationally focused tools and methods are needed in order to fully capitalize on the potentially rich information content of LiDAR data sets in the compilation of real-world forest vegetation inventories.

Acknowledgements

The authors wish to thank the two anonymous reviewers of our original manuscript for their thoughtful comments and suggestions for improving this paper.

References

Coops, N.C., Hilker, T., Wulder, M.A., St-Onge, B., Newnham, G., Siggins, A., Trofymow, J.A. 2007. Estimating canopy structure of Doulgas-fir forest stands from discrete-return LiDAR. *Trees-Structure and Function*, 21, 295-310.

Hilker, T., Wulder, M.A., and Coops, N.C., 2008. Update of forest inventory data with lidar and high spatial resolution satellite imagery. *Canadian Journal of Remote Sensing*, 34, 5-12.

King, D.J., 2000. Airborne remote sensing in forestry: sensors, analysis, and applications. *The Forestry Chronicle*, 76, 859-876.

Leckie, D.G. and Gillis, M.D., 1995. Forest inventory in Canada with emphasis on map production. *The Forestry Chronicle*, 71, 74-88.

Lim, K., Treitz, P., Wulder, M. A., St-Onge, B., and Flood, M., 2003. Lidar remote sensing of forest structure. *Progress in Physical Geography,* 27, 88-106.

Magnussen, S. and Russo, G., 2012. Uncertainty in photo-interpreted forest inventory variables and effects on estimates of error in Canada's National Forest Inventory. *The Forestry Chronicle,* 88, 439-447.

Means, J.E., Acker, S.A., Fitt, B.J., Renslow, M., Emerson, L., and Hendrix, C.J., 2000. Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering & Remote Sensing*, 66, 1367-1371.

Næsset, E., 2004. Practical large-scale forest stand inventory using small-footprint airborne scanning laser. *Scananavian Journal of Forest Research,* 19, 164-179.

Pinto, F., Rouillard, D., Sobze, J.M., Ter-Mikaelian, M., 2007. Validating tree species composition in forest resource inventory for Nipissing Forest, Ontario, Canada. *The Forestry Chronicle,* 83, 247-251.

Special Committee on Timber Supply, 2012. Growing Fibre, Growing Value. *Report to the Legislative Assembly of the Province of British Columbia*, 60 pp.

Treitz, P., Lim, K., Woods, M., Pitt, M., Nesbitt, D., and Etheridge, D., 2012. LiDAR sampling density for forest resource inventories in Ontario, Canada. *Remote Sensing*, 4, 830-848.

Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D., and Treitz, P., 2011. Operational implementation of a LiDAR inventory in Boreal Ontario. *The Forestry Chronicle,* 87, 512-528.

Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T., and White, J.C., 2008. The role of LiDAR in sustainable forest management. *The Forestry Chronicle,* 84, 807-826.

Zhao, K., Popescu, S., Meng, X., Pang, Y., and Agca, M., 2011. Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment,* 115, 1978-1996.

Vegetation classification in southern pine mixed hardwood forests using airborne scanning laser point data

Robert J. McGaughey[1], Stephen E. Reutebuch[2]

[1]USDA Forest Service, Pacific Northwest Research Station, bmcgaughey@fs.fed.us
[2]USDA Forest Service, Pacific Northwest Research Station, sreutebuch@fs.fed.us

Paper Number: SL2012-163

1. Introduction

Forests of the southeastern United States are dominated by a relatively small number of conifer species. However, many of these forests also have a hardwood component composed of a wide variety of species that are found in all canopy positions. The presence or absence of hardwood species and their position in the canopy often dictates management activities such as thinning or prescribed burning. In addition, the characteristics of the under- and mid-story layers, often dominated by hardwood species, are key factors when assessing suitable habitat for threatened and endangered species such as the Red Cockaded Woodpecker (*Picoides borealis*) (RCW), making information describing the hardwood component important to forest managers.

General classification of cover types using LIDAR data has been reported (Song et al. 2002, Brennan and Webster 2006) but most efforts focusing on the identification of individual species or species groups rely on some type of imagery to provide more complete spectral information for the study area. Brandtberg (2007) found that use of intensity data significantly improved LIDAR detection and classification of three leaf-off deciduous eastern species: oaks (*Quercus* spp.), red maple (*Acer rubrum* L.), and yellow poplar (*Liriodendron tulipifera* L.).

Our primary objective was to determine the proportion of hardwood species present in the canopy using only the LIDAR point data and derived products. However, the presence of several hardwood species that retain their foliage through the winter months complicated our analyses. We present two classification approaches. The first identifies areas containing hardwood and softwood (conifer) species (H/S) and the second identifies vegetation with foliage absent or present (FA/FP) at the time of the LIDAR data acquisition. The classification results were used to develop predictor variables for forest inventory models. The ability to incorporate the proportion of hardwood and softwood was important to the inventory as well as habitat assessments for the RCW.

1.1 Data

1.2 Study area

This study was conducted on the Savannah River Site (SRS). SRS is a National Environmental Research Park covering 80,267 ha (198,344 acres) located in the southeastern coastal area of the United States in west central South Carolina. In partnership with the Department of Energy (DOE), the USDA Forest Service's Savannah River Forest Station manages nearly 73,653 ha (182,000 acres) of commercial forest and more than 4,856 ha (12,000 acres) of non-forest land for a variety of natural resources.

Forests of the area are about 69% pine and 31% hardwood or mixed pine-hardwood. Dominant pine species include longleaf (*Pinus palustris* Mill.) and loblolly (*P. taeda* L.) pine and common hardwood species include various oaks, yellow poplar, blackgum (*Nyssa sylvatica* Marsh.), sweetgum (*Liquidambar styraciflua* L.), red maple, hickories (*Carya* spp.), and hollies (*Ilex*

spp.). Bottomland hardwood forests are found along SRS streams and on the "islands" or "ridges" of the Savannah River swamp. In these forests, typical canopy species include water oak (*Q. nigra* L.), laurel oak (*Q. laurifolia* Michx.), sweetgum, elms (*Ulmus alata* Michx. and *U. Americana* L.), red maple, and yellow poplar. Swamp forests are common along the western boundary of the site, adjacent to the Savannah River. Baldcypress (*Taxodium distichum* L.) and water tupelo (*Nyssa aquatic* L.) are common in these forests. Within the forests of SRS, not all hardwood species are deciduous. SRS has a mixture of hardwood species that are evergreen, tardily deciduous, or that retain desiccated leaves through winter.

1.3 Laser data

LIDAR data were acquired for SRS in the spring of 2009 when deciduous trees were in leaf-off condition. Data were acquired using two Leica ALS50-II laser scanners (designated as sensor 46 and sensor 77 by the data provider) mounted in separate fixed-wing aircraft and operating during the same time period. The average overall pulse density was approximately 10 pulses/m^2. The total area covered by the acquisition was approximately 119,000 ha (294,055 acres). Acquisition specifications for the data are shown in Table 1.

Table 1. Flight parameters and scanning system settings.

Flying height above ground (planned)	1432 m
Scan angle (flown)	±10°
Scan angle (delivered)*	±8°
Average scanning swath width (flown)	505 m
Swath overlap (flown)	62.5 percent
Scan frequency	58 Hz
Pulse rate	150 kHz
Beam divergence	0.22 mRad

*Returns from the outer 2 degrees of each scan were deleted prior to delivery. This reduced the scan angle, swath width, and overlap.

1.4 Field measurements

Tree measurements were collected on 194 plots in the spring of 2009. The plot protocol used a nested set of circular, fixed area plots to characterize all trees with diameter at breast height (DBH) larger than 2.54 cm (1 inch). The basic plots were 0.04 ha (0.1 acre) unless there were fewer than 8 dominant or co-dominant trees present on the plot. For these sparse stands, the plot size was increased to 0.081 ha (0.2 acre). On the basic plot the following measurements were taken for live and dead trees with DBH larger than 7.62 cm (3 inches): species, DBH, height, crown base height, and crown class. A smaller 0.004 ha (0.01 acre) plot, nested within the basic plot was used to collect detailed information for the smaller trees. On this plot the same measurements were taken for live and dead trees with DBH larger than 2.54 cm (1 inch) and less than or equal to 7.62 cm (3 inches). Smaller trees on the basic plot but outside the smaller 0.004 ha (0.01 acre) plot were tallied by species and size class (2.54 cm (1 inch) <= DBH < 5.08 cm (2 inch) and 5.08 cm (2 inch) <= DBH < 7.62 cm (3 inch)). For this study data were summarized by species, hardwood/conifer classes, live/dead, and foliage absent/present conditions.

A separate crew collected plot locations using dual-frequency, survey-grade GPS receivers (JAVAD Maxor. At least 600 positions were recorded for each plot center (10 minute occupation with 1-second epochs). Position data were post-processed using a continuously operating reference station (CORS) located close to the study site.

1.5 Photo measurements

Twenty five additional 0.04 ha (0.1 acre) plots were established using aerial photography acquired in leaf-off conditions. The 25 photo plots were located in swamp areas where the dominant species were baldcypress and water tupelo. Both species were without foliage at the time of the acquisition. Locations for these plots were digitized directly from the digital orthophotos in a GIS.

2. Methods

2.1 Intensity adjustment

Differences between the intensity values for the two sensors are visible in an image (7.62 m (25 feet) square pixel) produced using the first returns (Figure 1: left image). Intensity values for first returns were compared using 4,419 5- by 5-meter samples distributed evenly (10m spacing) along the area covered by the two sensors (red line in Figure 1). These samples were extracted from the overall point cloud and descriptive statistics were computed for the intensity values in each sample (Table 2). Overall, sensor 46 recorded higher values than sensor 77 over the same target area. In general, the difference in the intensity values recorded by the two sensors (Figure 2) is similar to the difference one would observe in photographs of the same area acquired using two different cameras or different exposure settings. Histogram matching is a relatively simple process that balances detector responses when dealing with images collected by different sensors or under different atmospheric conditions (Gonzalez and Woods 2007). A histogram matching procedure was implemented to adjust intensity values for first returns from sensor 46 relative to those from sensor 77.

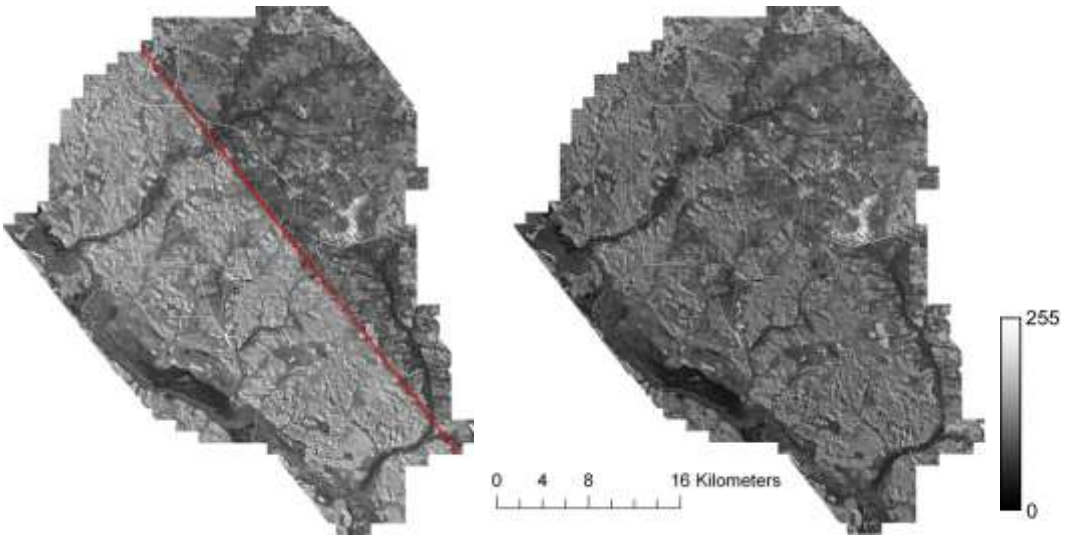


Figure 1. Intensity images created using original intensity values for first returns (left) and adjusted intensity values for first returns (right). The red line on the left image shows the center of the area covered by both LIDAR sensors. The NE portion was collected using sensor 77 and the SW portion using sensor 46.

2.2 Selection of training plots

For our supervised classification, we wanted to identify training plots that represented a “pure” condition. We selected a subset of field plots based on the mix of species present on each plot. We wanted plots with basal area composed of either all conifer or all hardwood species. From the 194 field plots, 19 conifer plots and 15 hardwood plots were identified.

Training plot selection was complicated by the presence of deciduous or tardily-deciduous hardwood species. When considering the presence or absence of foliage for trees on a plot, we

were able to identify only three plots with 98% or more of the basal area in species without foliage. Initially, we wanted all trees on a plot to have the same condition but found that none of the 194 plots met the criterion. To augment the data for the "foliage absent" condition, we used the 25 photo plots in addition to the three field plots. Thirty six plots had 98% or more of the basal area in species with foliage.

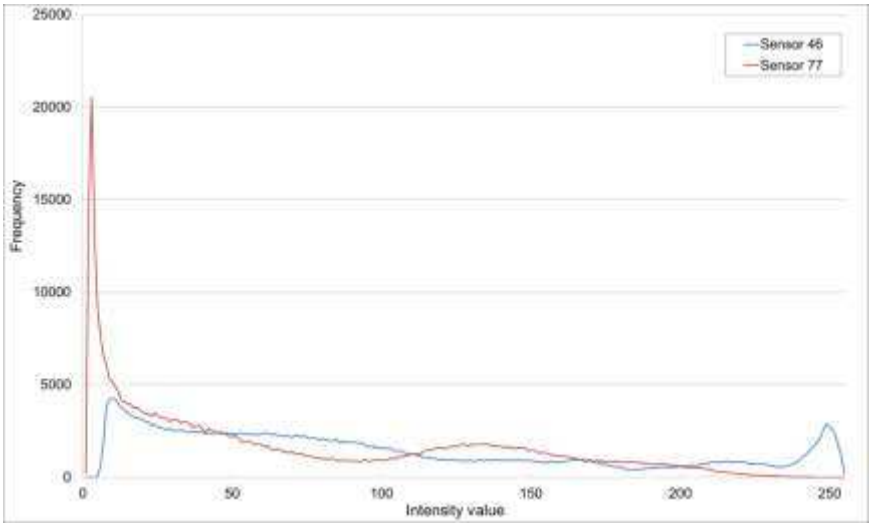


Figure 2. Distribution of first return intensity values for sensor 46 and sensor 77 for the portion of the acquisition covered by both sensors.

2.3 LIDAR-derived metrics

We extracted samples from the original point cloud and the first return data with adjusted intensity values using the location and size of each plot. We computed metrics for all of the plots using 2- by 2-m cells--resulting in 98 cells for the 0.04 ha (0.1 acre) plots and 197 cells for the 0.081 ha (0.2 acre) plots. For the H/C classification, there were a total of 4,189 cells with 2,071 cells representing pure hardwoods and 2,118 cell representing pure conifers. For the FA/FP classification, there were 6,660 cells with 2,831 representing vegetation without foliage and 3,829 representing vegetation with foliage.

Two types of metrics were computed for each 2- by 2-m cells; Intensity metrics using only first returns with adjusted intensity values within 2m of the canopy surface, and metrics computed using the height above ground for LIDAR returns above 2m. We also computed a pulse penetration metric based on first and last-of-many return surfaces but did not find it useful for classification.

To compute intensity metrics, a 0.5- by 0.5-meter resolution canopy surface model was used to isolate only first returns within 2m of the canopy surface. The returns in this sample were then compared to the LIDAR-derived ground model to eliminate all returns within 2m of the ground surface. The goal was to eliminate returns from understory vegetation, shrubs, and grasses that were not represented in field plot measurements. Finally descriptive metrics for the intensity values were computed using the remaining returns.

Height metrics were computed using the plot samples from the original return data. Return heights were computed by subtracting the return elevation from an elevation interpolated from the bare ground surface models using the XY location of the return. Metrics were computed using all returns above 2m. In addition to the standard set of metrics output by the FUSION GridMetrics program (McGaughey 2012), we also computed relative height percentile values by dividing height percentile values by the 95th percentile value. The relative percentiles are not

sensitive to the tree height on the plots and proved more useful in the classification process than the actual percentile values.

2.4 Classification

Metrics computed for the cells were used in the R statistical package (R Development Core Team 2010) along with the Rattle data mining GUI (Williams 2009) to conduct a supervised classification to identify conifer and hardwood vegetation and vegetation with foliage absent/present. We applied three classification methods: Random forest (Breiman 2001, Liaw and Wiener 2002), a simple decision tree approach (Breiman et al. 1984), and the adaptive boosting model (Friedman et al. 2000). The basic process used to build and apply the classification rules was:

- Conduct a Principle Components Analysis (PCA) to identify the components that explain the majority of the variation in the cell data,
- Evaluate correlations between components and LIDAR-derived metrics and select metrics most highly correlated with the components for use in building the classification rules,
- Conduct a classification using 70 percent of the cells to produce a set of 500 decision trees,
- Use the decision trees and data for the remaining 30 percent of the cells to evaluate the performance of the classifier,
- Apply the classification rules to cells for all plots at 2m cell resolution.
- Compute the proportion of each plot in each of the vegetation conditions at 20m cell resolution.

3. Results

3.1 Intensity adjustment

The adjusted intensity data were compared to the original data to evaluate the effectiveness of the intensity correction. An F-test indicated the original sample variances were not equal ($p = 1.98 * 10^{-32}$) and the variance of adjusted values were equal ($p = 0.36$) at the 5 percent level. The results of a t-test indicated the mean values of the samples were not equal in both the original data and the adjusted data. To further correct values for sensor 46, we applied a simple bias correction using the difference between the mean values for all samples. Table 2 shows the summary statistics for the original and adjusted intensity values and the right image in Figure 1 shows an image generated using the adjusted intensity for first returns.

Table 2. Summary statistics for first returns after intensity adjustment in 4,419 5- by 5-meter samples.

	Intensity values			
	Sensor 77	Sensor 46	Sensor 46 (after histogram matching)	Sensor 46 (after bias correction)
Mean	68	97	70	68
Standard deviation	35	41	35	35
Minimum	5	15	4	2
Maximum	217	249	210	207

3.2 Classification

PCA led us to select four variables to build the H/C classification rules (Table 3) and four variables to build the FA/FP classification (Table 4). Overall, the four variables in each set were not highly correlated and they provided values that describe the return intensity and the shape of the return height distribution.

Classification tools used 70% of the data to build rules and the remaining 30% to evaluate classification error. For the H/C classification, the overall error for the random forest method was 1.8%. For the FA/FP classification, overall error was 3.7%. The adaptive boosting model produced errors of 5.5% for the H/C and 5.8% for the FA/FP classifications. The simple decision tree produced errors of 10.8% for the H/C and 8.7% for the FA/FP classifications.

Table 3. Variables used in classification rules to differentiate hardwood and conifer species.

Variable	Description
Int.Mean	Mean intensity value for first returns close to the canopy surface (± 2m from canopy surface and height above ground greater than 2m)
Elev.RP30	30^{th} percentile (height) for all returns above 2m divided by the 95^{th} percentile for all returns above 2m
Percentage.all.returns.above.mean	Proportion of all returns above the mean height
Int.L3	Third L-moment for adjusted intensity values for first returns close to the canopy surface

Table 4. Variables used in classification rules to differentiate vegetation with and without foliage.

Variable	Description
Int.Mean	Mean intensity value for first returns close to the canopy surface (± 2m from canopy surface and height above ground greater than 2m)
Elev.RP40	40^{th} percentile for all returns above 2m divided by the 95^{th} percentile for all returns above 2m
Percentage.first.returns.above.mean	Proportion of first returns above the mean
Int.L.CV	L-moment coefficient of variation for adjusted intensity values for first returns close to the canopy surface

The classification rules were applied to the 2- by 2-m grid cells for all plots and the results summarized to produce “hardwood fraction” (HF) and “foliage absent fraction” (FAF) metrics. We compared these proportions to those computed using the basal area measured on field plots. Figure 3 compares the summarized cell classifications for the H/S classification and Figure 4 compares those for the FA/FP classification. In Figure 3 plots are separated based on the proportion of the basal area for hardwood species that retain their foliage through the winter. The red markers represent plots where more than five percent of the basal area is from species that retain their foliage and the green markers represent plots with less than five percent.

4. Discussion

A direct comparison of the H/S and FA/FP proportions is difficult since the plot data do not reflect the horizontal and vertical arrangement of trees and thus the proportion of a type visible from an aerial viewpoint may be different from the proportion measured on the ground. In reality, many of the hardwood species are shade tolerant and develop below a pine overstory. In many stands, the majority of the crowns associated with hardwood species would not be within 2m of the upper canopy surface and so returns from hardwood vegetation would not appear in the point-cloud data used to compute the intensity metrics.

For our data, the intensity values were not normalized to account for pulse strength, range, sensor gain, or atmospheric effects. While intensity normalization procedures have been described (Korpela 2008; Korpela et al. 2010), the process is not straight forward for the Leica ALS50-II sensors. This scanner provides automatic gain correction (AGC) to dynamically adjust the sensitivity of the detection circuitry to compensate for more or less reflective targets.

The gain settings and outgoing pulse strength were not available for use in a correction process so we adopted a more simplistic solution. The topography at SRS is generally mild with elevations ranging from 18 to 140 m. Flying height during the acquisition was consistent so the only range effects were due to the scan angle (±10 degrees as flown). Possible explanations for the different intensity values between the two sensors (Figure 2) include differences in laser power, detector sensitivity, manufacturer calibration, and hardware defects. Both instruments were configured identically and flown at the same height and during the same time period. Given the information available to us, we can offer no definitive explanation for the differences.

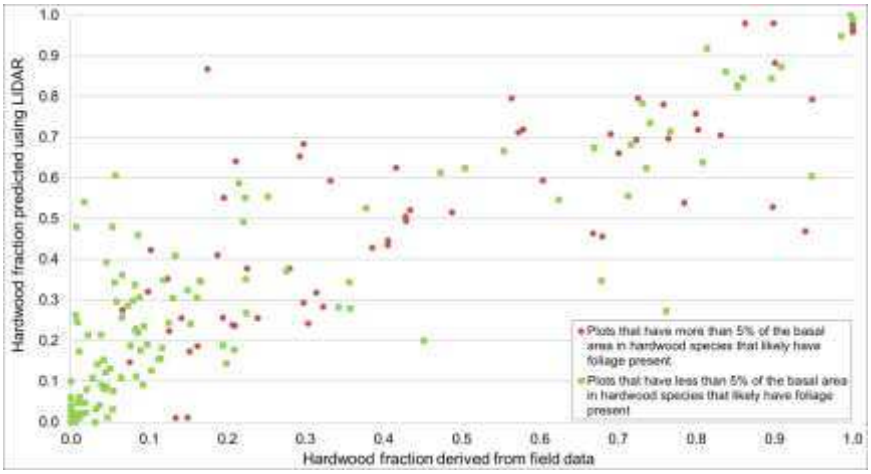


Figure 3. Scatterplot comparing hardwood/conifer fraction derived from field and LIDAR data.

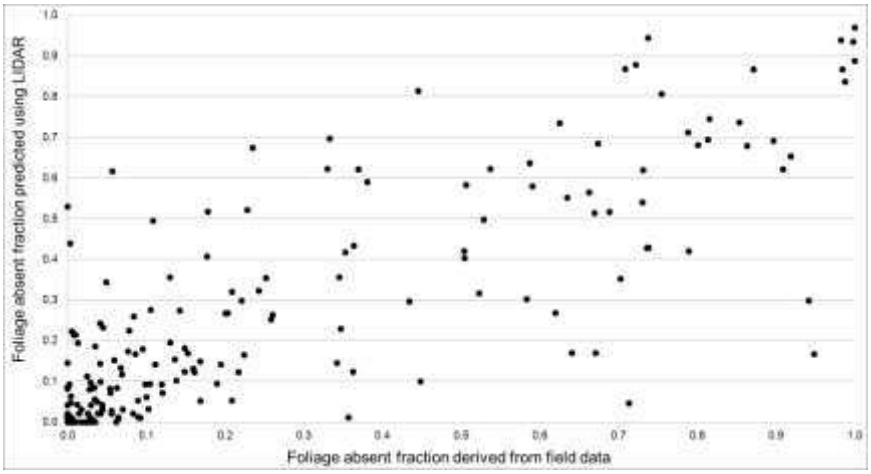


Figure 4. Scatterplot comparing foliage absent/present fraction derived from field and LIDAR data.

5. Conclusions

The results from our classification, proportions of hardwood/softwood and canopy with/without foliage, were used as predictor variables when modeling forest inventory parameters for SRS. The proportion of canopy with/without foliage was found to be a significant predictor variable for most models that predicted values for softwoods and hardwoods. Having the proportion available for use in regression modeling significantly improved the final models.

In general, our approach of building the classification rules using high-resolution data (2- by 2-meter pixels) and then summarizing the results at lower-resolution (20- by 20-meter pixels) produced results useful to managers. When we compare the results at both resolutions to ortho-

rectified aerial photos acquired in leaf-off conditions during the same year the LIDAR data were acquired, we see strong agreement (Figure 6). A similar approach would be useful for classifying live and dead trees in pure conifer types.

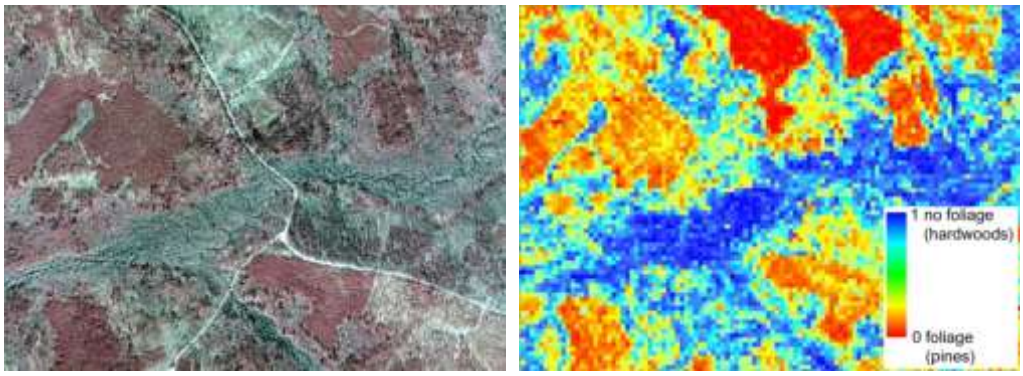


Figure 6. 2007 color infrared aerial photograph (left) and 2009 LIDAR proportion of canopy with/without foliage at 20m resolution (right).

Acknowledgements

Funding and support for this project was provided by the Department of Energy-Savannah River Operations Office through the U.S. Forest Service-Savannah River under Interagency Agreement DE-AI09-00SR22188. The authors would also like to thank the anonymous reviewer who made several suggestions that improved the paper.

References

Brandtberg, T. (2007). Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(5): 325–340.

Breiman, L., Friedman. J.H., Olshen, R.A., Stone, C.J. 1984. Classification and regression trees. Belmont, CA: Wadsworth.

Breiman, L. 2001. Random forests. *Machine Learning*. 45(1):5-32.

Brennan, R. and Webster, T.L. (2006). Object-oriented land cover classification of lidar-derived surfaces. *Canadian Journal of Remote Sensing*, 32(2):162-172.

Friedman, J., Hastie, T., and Tibshirani, R. 2000. Additive logistic regression: a statistical view of boosting. *Annals of Statistics*. 28(2): 337-374.

Gonzalez, Rafael C., Woods, Richard E. 2007. Digital Image Processing.Third Edition. Upper Saddle River, New Jersey: Prentice-Hall.

Korpela, I., 2008. Mapping of understory lichens with airborne discrete-return LiDAR data. *Remote Sensing of Environment*. 112(10):3891-3897.

Korpela, I., Orka, H.O., Hyyppa, J., Heikkinen, V., Tokola, T. 2010. Range and AGC normalization in airborne discrete-return LiDAR intensity data for forest canopies. *ISPRS Journal of Photogrammetry and Remote Sensing*. 64(4):369-379.

Liaw, A.; Wiener, M. 2002. Classification and regression by random forest. *R News*. 2(3):18-22.

McGaughey, R.J. 2012. FUSION/LDV: Software for LIDAR data analysis and visualization. Version 3.10. May 2012. http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf.

R Development Core Team. 2010. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*. Vienna, Austria. (http://www.R-project.org).

Song, J.-H., Han, S.-H., Yu, K., and Kim, Y.-I. (2002). Assessing the possibility of land-cover classification using LIDAR intensity data, *International Archives of Photogrammetry and Remote Sensing*, Graz, Austria, 2002, Vol. XXXIV, Part 3B, pp. 259-262.

Williams, G.J. 2009. Rattle: A data mining GUI for R. *The R Journal*. 1(2):45-55.

Characterizing grass, litter, and shrub fuels in longleaf pine forest pre- and post-fire using terrestrial lidar

Eric Rowell[1] & Carl Seielstad[1]

[1]National Center for Landscape Fire Analysis, College of Forestry & Conservation
The University of Montana, Missoula, MT 59812
eric.rowell@firecenter.umt.edu; carl.seielstad@firecenter.umt.edu

Paper Number: SL2012-166

Abstract

Empirically validating next-generation 3-dimensional fire behavior models requires detailed knowledge of fuel geometry in undisturbed fuel beds and researchers are turning to close-range remote sensing techniques such as terrestrial laser scanning (TLS) to characterize them. In fire-adapted longleaf pine forests, the problem is compounded by the spatial coincidence of shrub and herbaceous fuel elements with different fire behavior characteristics. In this paper, we derive a fuel height model (FHM) from TLS data to characterize grass, shrub, and litter fuels and use the FHM to assess structural differences between the fuel types. Co-incident pre- and post-fire scans are used to evaluate changes in normalized height distributions for grass, shrub, and hardwood plant communities at a grain of $0.5m^2$ on $16m^2$ plots. Fire-induced changes in fuel depth are distinct and logical between fuel types and perhaps related to the amount of fuel consumed. However, there are uncertainties due to differential shadowing of fuel elements pre- and post- fire and due to limitations in field data from undisturbed plots. Height data from field point sampling do not correspond directly with laser heights, probably due to difficulties in precisely matching point locations with scan data. Further, using a bare earth surface from pre-fire scans to normalize fuel heights produces sometimes large differences in the FHM compared with post-burn data, suggesting that occlusion from vegetation in even moderately loaded fuelbeds affects the accuracy of ground-normalized vegetation heights. Not surprisingly, the almost total consumption of grass and litter fuels is clearly evident in the post-burn data, while partially burned shrubs persist. Although comparison of pre- and post-fire scans produces intuitive results, it is not yet evident that the amount of fuel consumed can be estimated. TLS appears useful for describing the spatial heterogeneity of fuels at scales not possible with other techniques and remains a promising option for deriving 3-D fuel maps of larger landscapes in longleaf pine forest types. However, considerable research is still needed to prove up on the approach.

1. Introduction

The emergence of next-generation wildland fire behavior models that simulate fire propagation through 3-dimensional lattices at very fine grain has placed new demands on fuels data (Linn et al., 2002; Morvan and Dupuy, 2004; Mell et al., 2007; Pickett, 2008; Prince et al., 2010). The 3-D nature of fuels has long been recognized and tools were developed to translate them into 1-D values or qualitative indicators of fire behavior (Fahnestock 1970; Maxwell and Ward 1980; Anderson 1982; Burgan and Rothermel 1984); however, only recently have fuel descriptions and fire modeling advanced to the point where 3-D information can be used (Ottmar 2007; Parsons 2007).

Distributing fuels realistically within volumes is a difficult problem. Accurately measuring the 3-dimensional locations of fuels in the field is time consuming and not easily replicable. Additionally, abstracting these measurements to the attributes that models require is challenging. Nonetheless, recent

advances in laser scanning are facilitating more accurate and replicable 3-D rendering of fuels. For example, preliminary results from an experiment using a terrestrial laser scanner (TLS) to image Douglas-fir branches show that laser return density is highly correlated with fuel mass, while needles and small branches are distinct from larger branches and boles in laser intensity data (Seielstad et al., 2011). Moreover, the three-dimensional point clouds that result from laser scanning lend themselves naturally to production of the cell-based arrays (e.g., Hosoi and Omasa, 2007; Hiers et al., 2009) required by all of the current sub-grid models. Combined with the ability to scan on sub-centimeter grids at ranges to 1500 meters with distance and angular accuracies of 4-8mm, TLS offers considerable opportunity to advance fuels measurement science.

There are at least two primary routes for estimating biophysical characteristics from TLS data, 1) field measurements paired with laser scans of the natural environment, where multiple types, sizes, and arrangements of vegetation material are examined *in situ*, and 2) controlled laboratory experiments of individual vegetation elements such as shrubs, branches, or trees. With respect to wildland fuels, an example field based approach is described by Hiers *et al*. (2009). These authors integrated TLS data with thermal images of fire behavior in longleaf pine forest using local regression, in order to describe the types and scales of fuels and fire behavior. Loudermilk *et al* (2009) effectively used TLS to extract volumes of shrub fuels in a laboratory experiment, using a similar approach to the voxel method detailed by Hosoi and Omasa (2006) and Van der Zande *et al* (2006). Both experiments indicate that information regarding vertical arrangement and size of vegetation elements are useful for estimating volumes, plant area densities, and comparisons related to biomass. Beyond the two experiments described above, the application of TLS to surface fuel characterization and mapping is limited. Thus, drawing on research that focuses on estimations of fuel-like variables from the broader terrestrial lidar literature is essential. For example, the foliar and woody biomass estimates from Douglas-fir branches in a controlled setting demonstrated efficacy of using TLS (Seielstad *et al*. 2011) for discriminating fine fuels; however integration of individual branches to an entire tree is significantly more complex. Direct estimates of dry weight and lidar derived plant area density in wheat fields suggest that voxel-based canopy profiling may be useful for producing independent estimates of volumes and biomass in grass fuels (Hosoi and Omasa 2009). The latter study used a series of scans across a range of plant growth for $1m^2$ plots within a wheat field, where plant area is a function of voxel penetration and integration of leaf thickness over the total height of a wheat cluster. However, given the considerable uncertainties in characterizing specific fuels attributes such as mass by size-class or surface area to volume, the potential advantage of TLS for providing improved fuels data in the short term is in mapping the types and relative amounts and arrangements of vegetation materials across small landscapes (up to a few ha) with the purpose of assisting in validation of fire behavior modeling. The goal of such an effort is to generate rational, spatially-explicit estimates of fuels at fine grain (circa $0.25m^2$) for plots that will be burned using prescribed fire. Invariably, the landscapes being studied cannot be disturbed because they will be burned, thus, a remote sensing approach is warranted. Fire behavior model requirements for fuels data are fairly straightforward. A collection of cells within a model domain defines the geometry of the fuelbed, and each cell in the domain contains an estimate of bulk density of thermally-thin vegetation and packing ratio. The latter two terms require fuel mass by size class, volume occupied by the fuel, particle density, particle surface-area-to-volume (SAV) ratio, char fraction, specific heat, moisture content, and initial temperature. Realistically, density, SAV, char, specific heat, moisture, and temperature are estimated from laboratory measurements and field measurements at the time of burning. The TLS, then, is expected to describe the geometry of the fuel bed, and perhaps the type and relative amount of fuel in each cell. Specific attributes such as density, SAV, and mass by size class would still come from laboratory measurements.

In this study, we use TLS to image an intermix of litter, grass, and shrub fuels in a longleaf pine forest in Florida, USA, in order to characterize the relative amounts and arrangements of vegetation material pre- and post-fire. This research represents the basis for a larger goal of identifying the type and amount of

flammable materials within volumes in order to ultimately infer other fuel properties such as mass by size class and surface-area-to-volume ratio.

1.1 Objectives

The objectives of this research are threefold:

1. Derive a rational fuel height model (FHM).
2. Use the FHM to assess structural differences between shrub, grass, and litter fuel types.
3. Quantify post-fire changes to the FHM and relate them to fuel type.

The research is largely exploratory and inference from it is constrained by the amount and type of independent field measurements. Beyond point-intercept sampling of fuels, the study area is described by photography and optical estimation of fuel cover, type, and height. Biomass samples were taken adjacent to, but outside of the study area. We advance the concept of the Fuel Height Model due to its simplicity and similarity to the canopy height model (CHM) of airborne laser scanning.

2. Method

2.1 Study area

Data were collected at Eglin Air Force Base, FL, in February 2011. Eglin AFB is located on the panhandle of Florida, USA, originally a unit of the former Choctawhatchee National Forest; Eglin is an important resource in the management of longleaf pine ecosystems with 180,000 ha of longleaf pine sandhills and flatwoods. Eglin has one of the largest prescribed burning programs in the United States, with 40,000 ha burned annually, and is managed jointly by the United States Air Force and the Center for Environmental Management of Military Lands at Colorado State University. Plots with dimensions of 4m x 4m were established in two upland longleaf pine stands, which were subsequently burned with strip head fires ignited by a DAID ping pong ball system mounted on a rotor-wing aircraft. Plots were located in representative grass/shrub fuel types and instrumented with a variety of meteorological and thermal measuring devices prior to burning. For this experiment, we utilized three 4m x 4m plots containing a mix of grass, pine litter, and turkey oak shrub fuels. Fuels were sampled by point-intercept on a 0.5 meter grid from temporary scaffolding erected over each plot. At each sample point, the maximum height of each fuel type present was recorded (e.g., wire grass, other grasses, herbaceous, shrubs by species, litter cover and depth by type (pine and oak), count of woody debris, pine cones, and bare soil). Additionally, fuels were destructively sampled in ten 0.5 m^2 subplots per 4m x 4m plot outside of, but within 40 meters of the plots. The latter samples were oven-dried and weighed for bulk density. Plots were photographed at nadir from a height of 7 meters and photographed obliquely from each scan location pre- and post-fire.

Table 1. Plot Characteristic Summary

PlotID	Description	Mean Height (cm)	Height Range (cm)	Mean Load (Mg/ha)	Load Range (Mg/ha)
703E	Grass/Shrub	16.8 $\pm$ 14.8	1.0 – 53.0	4.7 $\pm$ 1.0	3.0 – 6.0
703W	Litter/Shrub	14.0 $\pm$ 12.6	3.0 – 57.0	5.1 $\pm$ 1.8	2.5 – 7.8
608NW	Forb/Grass	27.5 $\pm$ 15.9	3.0 – 75.0	4.8 $\pm$ 1.9	3.2 – 10.3
608SW	Shrub/Litter	20.2 $\pm$ 24.7	1.0 – 125.0	6.5 $\pm$ 2.6	2.2 – 11.1

2.2 TLS Data Collection and Processing

Laser scans were collected at each plot pre-and post-burn using an Optech ILRIS^{TM} 3_6D-HD instrument scanning at 10 kHz. The TLS was positioned to vantage points allowing for direct line of sight over the plot, three meters from plot edge and 2.7 meters above ground. Data were collected from three locations per plot, the northwest corner, northeast corner, and the center of the south side at a 6 mm scan resolution. Plot corners were referenced with a plastic cap on rebar. The northwest corner was double-capped for reference. Each scan was corrected to a control data set that was geo-spatially referenced to GPS data collected around each plot using Polyworks software suite (Innovmetric, Quebec, Canada). Scan to scan corrections were completed by selecting identifiable points within the control data and then further refining the scan correction using the automated align algorithm within Polyworks. The corner reference points were more easily identified in post-burn scans, therefore these points were used to reference adjacent scan scenes together. All laser scans were collected from the same locations both pre- and post-fire, pre-fire scans were aligned by using the location of the post-burn scan head as the primary control point. Matching of scans was highly dependent on the auto align algorithm in Polyworks, the absence of hard targets in the pre-fire scans made matching of scenes difficult.

2.3 Classifying Fuel point from Bare Ground, and Bare Earth DEM

Normalization of height distributions using a bare ground surface utilized a local minimum filter using a tessellated surface derived from the post-fire scans for each plot. The surface is derived using a triangulated grid; a 21 cm kernel is passed over the bare earth surface using the original point array from the scans decimated to a point every centimeter. The location of the easting and northing of the point is normalized to the minimum of the extents of the data and converted to an integer representing the corresponding pixel location of the point. The minimum of the local area encompassed by the kernel represents the new value of the surface. This method was used for both pre- and post- fire merged scans for all plots to assess relative accuracy for pre-fire derived bare earth surfaces.

2.3.1 Normalizing the Height Distributions

Heights were normalized to bare earth using a DEM generated from the post-fire scans to produce the FHM. Percentile heights were also generated from the FHM. Lastly, heights were binned in 1-cm increments and each bin was normalized by the total number of laser points in the cell. The latter approach was adopted following Ku *et al*. (2012), who note that height bin normalization can be more easily interpreted than percentile normalization, an important consideration given that our work is largely exploratory and that the actual heights of fuels are important to fire behavior in shrub/grass fuels.

2.3.2 Fishnet and Classifying Plots

Each 0.25m^2 grid cell was classified as shrub, grass, forb, or litter using a combination of point intercept data, nadir photography, and oblique photography. Cells were classified from the nadir photography by dominant cover and compared to point intercept data. If there was agreement between the two, the cell was classed correspondingly. If shrub was identified in the point intercept, the cell was classified as shrub. Where there was disagreement between classifications, oblique photography was consulted and a final call was made. Where ambiguity existed, the point intercept data prevailed. It is worth noting that there were very few homogenous cells in the non-shrub classes and discriminating between grass and forb was sometimes difficult. The final classification of non-shrub, then, represents dominant cover in the cell.

2.3.3 Extracting Height Metrics

Metrics were extracted to characterize height distributions for all plots and grid cells (n=180) excluding 0.25m^2 grid cells adjacent to the steel rebar in each corner of the plot. For each grid cell, histograms at 1cm bin size were analyzed to extract the point where the histogram becomes either asymptotic or drops below a threshold of 500 after the maximum frequency of the histogram occurs. For each grid cell, the inflection point, maximum frequency value, and ratio of points above and below the inflection point are recorded. Percentile, mean, and standard deviation of laser heights were also generated per grid cell, though not presented as no meaningful relationships could be derived with heights from the point intercept data.

3. Result

3.1 *The Fuel Height Model*

Table 2. Deviation of Pre- and Post-Fire Bare Earth Surfaces by Fuel Type

	Fuel Type							
	Forb		**Grass**		**Litter**		**Shrub**	
PlotID	Mean (cm)	Range (cm)	Mean (cm)	Range (cm)	Mean (cm)	Range (cm)	Mean (cm)	Range (cm)
703E	-	-	-0.6 ± 2.7	-5.7 – 12.6	-0.8 ± 3.9	-6.4 – 12.1	-2.1 ± 1.3	-4.8 – 2.6
608NW	-1.3 ± 2.3	-5.8 – 9.2	-1.6 ± 2.8	-6.9 – 6.4	-1.5 ± 2.5	-5.8 – 9.5	-1.4 ± 2.3	-6.4 – 7.6
608SW	-	-	-	-	0.2 ± 1.6	-4.8 – 6.6	0.1 ± 1.8	-6.0 – 9.3

We start from the assumption that post-fire laser height data closely approximate bare earth conditions because fire-induced consumption of grasses, forbs, and perched litter was near-complete and little litter and duff remained due to the fact that the stands had been burned every second year for the past decade. Coarse woody debris was rare. A comparison of pre-fire bare earth surfaces with post-fire surfaces, shows modest differences (table 2), averaging -1.0 cm across all plots. Between plot difference is more substantial than between fuel type difference. A negative difference means that the bare earth surface from the pre-burn scan is lower than the post burn bare earth surface, a non-intuitive result addressed below in Discussion. The best agreement occurs in plot 608aSW which is predominantly upright oak shrubs and flat oak litter. In each plot, the distribution of deviations centers near zero, is symmetrical, and is leptokurtic.

3.2 *Structure Differences between Fuel Types*

Although there is considerable variability in fuel height distributions within fuel types due to variations in the amount, type, and arrangement of material in each, the fuel types have characteristic appearances (figure 1). Grass-dominated cells tend to have a low peak (table 3) representing the bulk of leaves near the ground with an upright tail depicting narrow stems topped by inflorescence. Litter resembles grass, with a typically lower center of mass. Forb-dominated cells are multi-modal, with considerable reflections throughout the height distribution. Shrubs resemble grass, but have more material high in the distribution and are often taller.

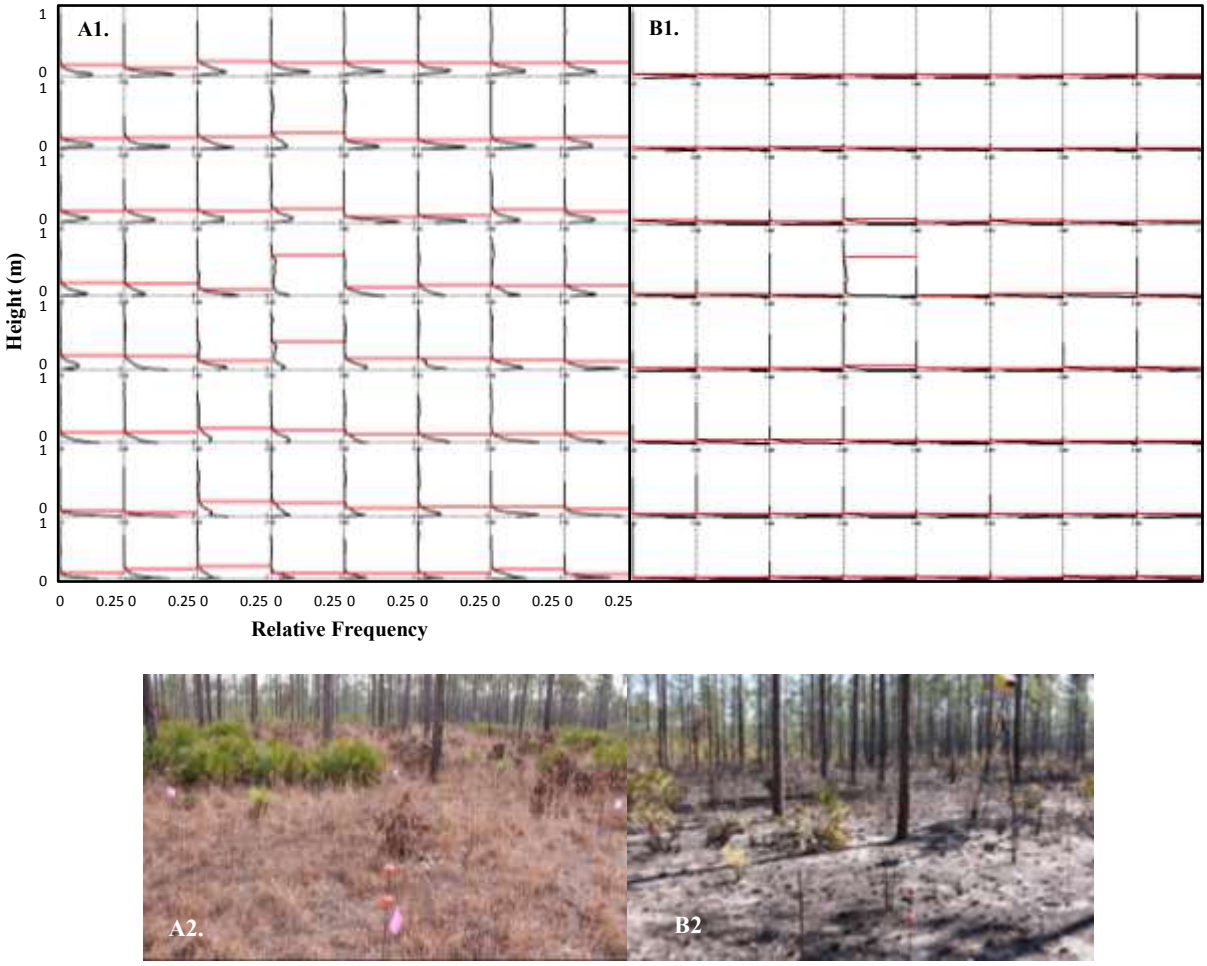


Figure 1: Histograms for site 703CE organized by grid cell layout for pre-burn (A1,picture A2.) and post-burn (B1, picture B2.). The red line across each histogram represents the inflection point for this grass dominated site.

Table 3. Distribution Metrics Derived from Histogram Analysis

Description	n	Inflection Height (cm)	Peak Histogram Height (cm)	Points Above Inflection (%)	Points Below Inflection (%)
Pre-Fire					
Forb	32	20.8 ± 17.2	12.1 ± 15.0	0.2 ± 0.7	99.8 ± 0.4
Grass	30	16.4 ± 3.6	3.0 ± 2.6	6.7 ± 1.8	93.3 ± 0.9
Litter	51	19.8 ± 12.8	11.7 ± 1.6	5.5 ± 1.9	94.5 ± 10.3
Shrub	67	35.5 ± 17.9	18.4 ± 11.2	4.1 ± 2.6	95.9 ± 0.05
Post-Fire					
Forb	32	4.3 ± 1.6	0.5 ± 0.4	2.0 ± 0.07	98.0 ± 0.7
Grass	30	6.7 ± 1.1	1.9 ± 0.1	0.3 ± 0.1	99.7 ± 0.9
Litter	51	6.2 ± 3.1	1.7 ± 0.2	1.0 ± 0.3	99.0 ± 0.3
Shrub	67	6.9 ± 7.0	1.2 ± 2.8	19.0 ± 4.3	81.0 ± 4.3

Multiple comparisons analysis indicates that across inflection point and peak frequency there are differences between shrub type fuels in comparison to grass and litter fuel types ($p < 0.005$). Grass and

litter fuel inflection points do not significantly differ, however both fuel types differ in regards to the peak frequency and also are divergent from the mean peak of the shrub fuel type ($p < 0.005$). The forb fuel class only differs from the other classes when comparing points above and below the inflection point ($p < 0.005$); all other classes are indistinguishable in this regard.

3.3 *Post-fire Changes to Fuelbeds*

Post fire characteristics (table 3) demonstrate differences of the four fuel types in terms of the pre-fire scan data. Expectedly, variability in height distributions becomes negligible between the forb, litter, and grass fuel types in the post-fire environment. The shrub fuel class stands as the exception, where hardwood retention post fire accounts for 19% of laser points that reside above the inflection point as compared to 0.3-2.0% for the forb, grass, and litter classes. Multiple comparison analysis supports this result, as the shrub class is the only significantly different fuel type for ratio of points above and below the inflection ($p<0.005$). Inflection and peak histogram values for all fuel type are statistically the same, where inflection and peak frequency both lower in height to characterize bare ground. In the frequently burned environment of Eglin AFB, coarse woody debris is sparse, though in plot 608ANW there is evidence of slightly higher inflection heights where post-fire inflection height average 6 cm higher than non-shrub fuel types in the plot which is attributed to a single log lying across the plot. Grid cells across forb, grass, and litter fuel types exhibited small proportions of points above the inflection point motivated by retention of emergent turkey oak hardwood previously masked by grass and forb cover.

4. Discussion

As previously mentioned, TLS research of surface fuels is limited and studies previously conducted have attempted to quantify elements of fuel beds in regards to optimal grain size (Hiers *et al*. 2009) and quantifying fuel biomass in a controlled setting (Loudermilk *et al*. 2009). Our research seeks a framework that ultimately distinguishes fuel types and provides a mechanism to distribute biomass estimates across fuel types and the landscape, this paper is an initial exploratory study to begin developing the methods to meet these objectives. The fuel height models derived in this work are rational given their similarities to what is observed and produce readily identifiable contrasts that allow for classification of fuel type and potential location of the bulk of biomass. Characterization of fuel types and distributions demonstrate that the forb fuel type engenders characteristics associated with both shrub and grass fuel classes in the pre-fire data. Aspects of the fuels class confounding the results most likely stem from the need to generalize fuel classes in highly variable fuel beds. Using majority rules for fuel typing, elements such as the litter fuel type in fact contain variable stocking of both grass and forb. Inevitably, generalizations of fuel bed characterization will always be needed at the cost of accuracy. The work described in this paper is tied to a larger set of integrated measurements and points to some significant shortcomings. For instance, the point intercept data collected does not accurately describe the full variability in a grid cell for both fuel type and height. Traditionally, these point intercept data are linked to height and cover estimates that are used to predict biomass from destructive sampling in similar conditions, though in the case of this experiment these relationships have been poor. Collecting validation data comparable to laser scanned data is difficult, requiring dense sampling protocols to adequately quantify fuel height, type, and amount at the grain size used in next generation fire behavior modeling. Selection of the 0.5 meter cell size is designed to represent the spatial heterogeneity implicit

at fine scales in the southeastern United States long-leaf pine ecosystem (Hiers *et al*. 2009), but this grain size nears the bottom limit of what fire behavior models realistically use. Selection of fuel cell sizes that reflect actual fuels variation may better mitigate the issue with splitting fuel types between cells. Noting the similarities of post-fire removal of aboveground data points, assessing multiple grain sizes to address intermixing of fuel types, or generalization of fuels into shrub and other fuels may be an alternative. The retention of aboveground data points in the shrub fuel type post-fire begs the question as to whether there will always be a need for post-fire data to identify this fuel type. Compelling findings of this work suggest that accurate and consistent bare earth surface generation can be derived from both the pre- and post-burn data and from these surfaces a FHM can be developed to extract meaningful metrics representative of fuel characteristics; we feel this is a significant finding. The lower pre-fire surfaces are potentially a result of the auto-aligning process used to refine the matching of adjacent scans. Potentially, the resulting change of inflection point and peak histogram frequency from fire may be used in a predictive sense to estimate mass loss due to fuels consumption. In the future, we suggest a concept for defining fuel cells using combinations of peak frequency and inflection point to weight and attribute fuel cells used in advanced fire behavior models. We envision segmenting the fuel cells using breakpoints in the FHM to describe fuel geometry, thus allowing for the distribution of bulk density estimates into partitions representative to where fuel actually exists.

5. Conclusion

Presently, fuels characterization for fine-grained modeling purposes relies on destructive sampling outside of the study domain to develop probability distributions which are subsequently used to model fuels within the areas to be burned. Using TLS to inform this modeling in a spatially explicit way seems logical in light of the behavior of the FHM in the fuel beds described in this paper. There is considerable further investigation required to prove up on the approach, much of it under way. For example, TLS data are being compared with fuel attributes from destructively sampled fuel beds to develop predictive models relating height distributions to fuel type, amount, and arrangement. The limits to TLS for mapping and quantifying fuels at very fine grain are still unknown.

Acknowledgements

Many thanks to: Joe O'Brien and Ben Hornsby of the USFS Southern Research Station in Athens, GA, for access to field plots and data; the staff of Jackson Guard at Eglin AFB for access and for planning and implementing the burns; RxCadre for organizing the field campaign, and the National Center for Landscape Fire Analysis at The University of Montana, for funding.

References

Anderson HE (1984) Aids to determining fuel models for estimating fire behavior. Gen. Tech. Rep. INT-122. Ogden, Utah: U.S.Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 22p.

Burgan, R. E.; Rothermel, R. C. 1984. BEHAVE: fire behavior prediction and fuel modeling system FUEL subsystem. General Technical Report INT-167. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Range and Experiment Station.

Fahnestock, G. R. 1970. Two keys for appraising forest fire fuels. Research Paper PNW-99. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station

Hiers, J.K., J.J. O'Brien, R.J. Mitchell, J.M. Grego, and E.L. Loudermilk. 2009. The wildland fuel cell concept: an approach to characterized fine-scale variation in fuels and fire in frequently burned longleaf pine forests. International Journal of Wildland Fire, 18: 315-325.

Hosoi, F. and Omasa, K. 2008. Estimating vertical plant area density profile and growth parameters of a wheat canopy at different growth stages using three-dimensional portable lidar imaging. ISPRS Journal of Photogrammetry and Remote Sensing, 64: 151-158.

Hosoi, F. and Omasa, K. 2007. Factors contributing to accuracy in the estimation woody leaf area density profiles using portable lidar imaging. Journal of Experimental Botany, 58(12): 3464-3473.

Hosoi, F., Omasa, K., 2006. Voxel-based 3-D modeling of individual trees for estimating leaf area density using high-resolution portable scanning LiDAR.IEEE Transactions on Geoscience and Remote Sensing 44, 3610–3618.

Ku, N-W, Popescu, S., Ansley, R., Perotto-Baldivieso, H., and Filippi, A. 2012. Assessment of available rangeland woody plant biomass with a terrestrial lidar system. Photogrammetric Engineering and Remote Sensing, 78(4): 349-361.

Linn R, Reisner J, Colman JJ,Winterkamp J (2002) Studying wildfire behavior using FIRETEC. International Journal ofWildland Fire 11, 233–246.doi:10.1071/WF02007.

Loudermilk, E.L., Hiers, J.K., I'Brien, J.J., Mitchelm R.J., Singhania, A., Fernandez. J.C., Cropper, W.P., and Slatton, K.C. 2009. Ground-based LIDAR: a novel approach to quantify fin-scale fuelbed characteristics. International Journal of Wildland Fire. 18, pp 676-685.

Maxwell WG, Ward FR (1980) Photo series for quantifying natural forest residues in common vegetation types of the Pacific Northwest. Gen. Tech. Rep. PNW-GTR-105. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station

Mell W, Jenkins MA (2007a) A physics based approach to modeling grassland fires. Int. J. Wildand Fire 16: 1-22

Morvan, D. and J. L. Dupuy, "Modeling the Propagation of a Wildfire through a Mediterranean Shrub Using a Multiphase Formulation," Combustion and Flame, 138, 199-210 (2004).

Ottmar RD, Sandberg DV, Riccardi CL, Prichard SJ (2007) An overview of the Fuel Characteristic Classification System—quantifying, classifying, and creating fuelbeds for resource planning. Canadian Journal of Forest Research 37, 2383–2393.

Parsons RA (2007) Spatial variability in forest fuels: simulation modeling and effects on fire behavior. Ph.D. dissertation, University of Montana, Missoula, MT. 255 p.

Pickett, B. M., Isackson, R. Wunder, T. H. Fletcher, B. W. Butler, and D. R. Weise, "Flame Interactions and Burning Characteristics of Two Live Leaf Samples," International Journal of Wildland Fire, 18, 865–874 (2009).

Prince, D., B. Andersen, W. Cole, M. Dennis, T. H. Fletcher, "Modeling a Burning Shrub with and without Wind using a Semi-empirical Model," presented at the International Association of Wildland Fire 3rd Fire Behavior and Fuels Conference, Spokane, Washington (October 25-29, 2010). (paper in progress).

Seielstad, C., Stonesifer, C., Rowell, E., and Queen L. 2011. Deriving fuel mass by size class in Douglas-fir (Pseudotsuga menziesii) using terrestrial laser scanning, Remote Sensing, 3: 1691-1709.

Van der Zande, D., Hoet, W., Jonckheere, I., Van Aardt, J., and Coppin, P., 2006. Influence of measurement set-up of ground-based LiDAR for derivation of tree structure. Agricultural and Forest Meteorology, 141:147-160.

Dynamic, LiDAR-based assessment of lighting conditions in Pacific Northwest forests

Demetrios Gatziolis[1]

USDA Forest Service, PNW Research Station, e-mail:dgatziolis@fs.fed.us

Paper Number: SL2012-170

Abstract: Light Detection and Ranging (LiDAR) technologies employed over forested landscapes provide a detailed representation of dominant objects, typically tree crowns and ground surfaces. In this study, the actual lighting regime within a forest stand is derived from LiDAR data in a two-stage process. First, a LiDAR data-derived ground surface raster is used to determine whether physiographic conditions around the targeted area affect direct sun illumination. The second stage operates on the voxel domain, which partitions space in discrete cubical elements and labels them either filled, if they contain at least one laser return, or empty otherwise. A buffer around the area of interest ensures that shadowing induced by surrounding vegetation is considered in the computations. Ray tracing through the voxel space along the trajectory that connects the sun and each facet of each filled voxel classifies the facets as either sunlit or in shadow. The process can be repeated for additional sun locations or time intervals from sunrise to sunset for a given day. Ultimately, the daily cumulative or time-interval-specific amount of direct sunlight during a cloudless day can be computed. Adequate hardware resources are critical for reasonable computing times over large areas or fine time intervals. The methodology described above was evaluated using precisely georeferenced and timed field observations of ground illumination conditions in forest stands in eastern Oregon. Remarkable agreement between recorded and derived lighting conditions was observed. This methodology for computing lighting regime wherever high-density LiDAR data are available has various potential applications including contributions to forest growth and yield models, assessment of regeneration potential, stream shading, and in support of management and tactical decision making.

1. Introduction

Sunlight is a major environmental and ecological factor. Within forest stands, sunlight, or simply light, availability is affected by the amount and spatial distribution of above-ground vegetation (Parent and Messier, 1996), topography, and by diurnal and seasonal cycles in the sun trajectory. Light availability is known to limit the survival and growth of many forest species (Comeau *et al.* 1993; Nicotra *et al*. 1999), and to affect agroforestry system practices (Bellow and Nair 2003), the dispersal of invasive species (Parendes and Jones, 2000), stream temperature (Chen *et al*. 1998), and many other phenomena and activities. Considering the high variability of light availability in space and time observed in field conditions, the explicit, distributed assessment of light conditions is regarded as an ill-posed task primarily due to the tremendous, and seldom met data requirements on the structural elements of a forest stand (Brunner, 1998). To overcome these limitations, various simplifying assumptions are made concerning the main structural elements of a forest stand, and light availability is assessed via modeling (e.g. MacFarlane *et al*. 2003). Tree crowns in the models are represented by simple geometrical shapes such as cones, cylinders, or ellipsoids. Model predictions routinely serve as input to growth and yield models or in support of forest management decisions. Field measurements of instantaneous light availability necessary for model development are typically obtained by processing hemispherical photos, or by using quantum sensors during overcast conditions. All options add substantial logistical cost to an already complex effort.

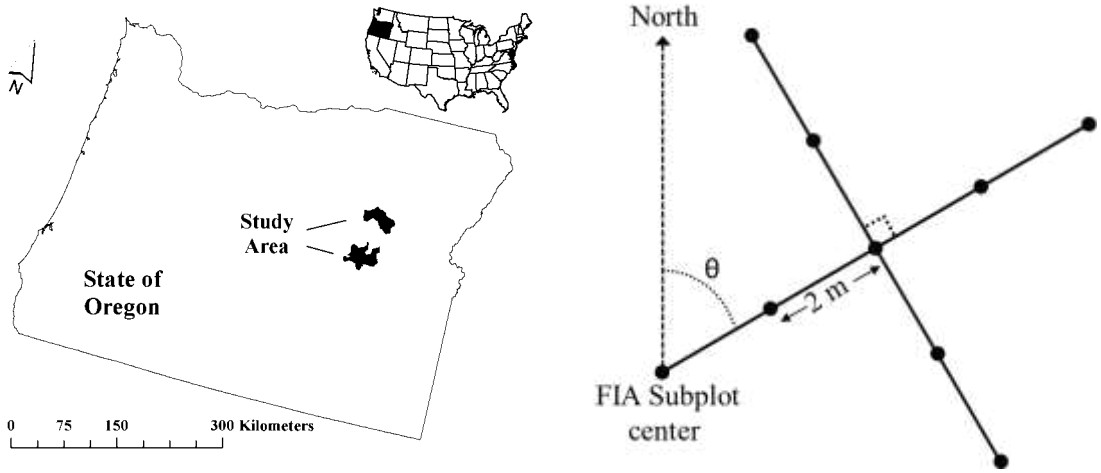
Accurate assessment of temporal sun illumination regimes within forests might be attainable where Light Detection and Ranging (LiDAR) data are available. Assuming adequate density in the return point cloud, renditions of the LiDAR data in three-dimensions (3D) presented to humans are perceived as detailed representations of tree crowns and of the ground, the dominant structures in a forested scene. If LiDAR data could be consistently converted to a 3D representation of vegetation objects, the interaction between these objects and the light emitted from a source positioned somewhere within or outside the scene could be calculated analytically. A simplified variant of this approach is emerging with increasing frequency in forest applications. It involves a highest-return raster surface generated from the LiDAR data that is processed with an implementation of the hillshade algorithm (eg. Burrough and McDonell, 1998) for a specified sun azimuth and elevation angles to determine the sunlit cells and those in shadow. While this approach might be a useful approximation of lighting regimes when the sun is at a high elevation angle, unfortunately it is biased and favors cells in shadow. The bias is introduced primarily by the consistent exaggeration of the volume occupied by tree crowns and also by the inability of the raster representation, sometimes known as a 2.5D rendering, to consider illumination of vegetation and ground portions that are overtopped by vegetation. This study introduces an alternative approach potentially capable of providing accurate illumination estimates. Its performance is evaluated using actual light availability observations in eastern Oregon.

2. Methods

Voxel-based modeled and LiDAR-based representations of ground surfaces and vegetation are intersected, or traversed, by rays that mimic solar illumination. A series of simulations for different sun positions are performed and the resulting illumination regimes are summarized.

2.1 Study area and field data

The approximately 190,000 ha study area is located in the Malheur National Forest in eastern Oregon, USA, where elevation ranges from 1,250 to 2,600 m (Figure 1a). Dominant species include Ponderosa pine (*Pinus ponderosa*), Douglas-fir (*Pseudotsuga menziesii*), western larch (*Larix occidentalis*), grand fir (*Abies grandis*), and lodgepole pine (*Pinus contorta*). The Forest exhibits a variety of physiographic conditions and the majority of forest stands feature open canopies with medium, between 30 and 60 percent, canopy cover. The study area contains 186 Forest Inventory and Analysis (FIA) Program plots with more than 5,000 tallied trees. FIA uses



a. b.

Figure 1: a. Study area in the Malheur National Forest, and b. Field protocol for solar illumination observations

a clustered design with each plot comprising 4 subplots. Solar illumination conditions were recorded in 54 FIA subplots in the summer of 2010. Randomly oriented transects were established in each subplot arranged as shown in Figure 1b and illumination conditions were recorded at two meter intervals from the edge of each transect. Each observation was timed and assigned a code; 0 if in shadow, 2 if sunlit, and 1 if very close to the boundary between sunlit and shadowed areas. An index value, computed as the mean of all observations per subplot, quantified the solar illumination conditions. Coordinates of submeter precision were obtained at each subplot center via post-processed differential Global Positioning System data.

2.2 LiDAR data

Discrete return, airborne LiDAR data were acquired over the study area in 2008 and 2009 with two instruments; a Leica ALS50 Phase II and a Leica ALS60. Both operated with pulse frequency of 105 kHz and are capable of recording up to 4 returns per pulse. An average 55 percent scanning swath sidelap, mean instrument-target range of 920 m, mirror scan frequency of 52, and return filtering to a maximum scan angle of 14 degrees yielded a mean density of 9 returns per square meter. The nominal footprint diameter and spacing were 28 and 68 cm respectively. A 0.91-m, LiDAR-derived Digital Terrain Model (DTM), was provided by the data vendor.

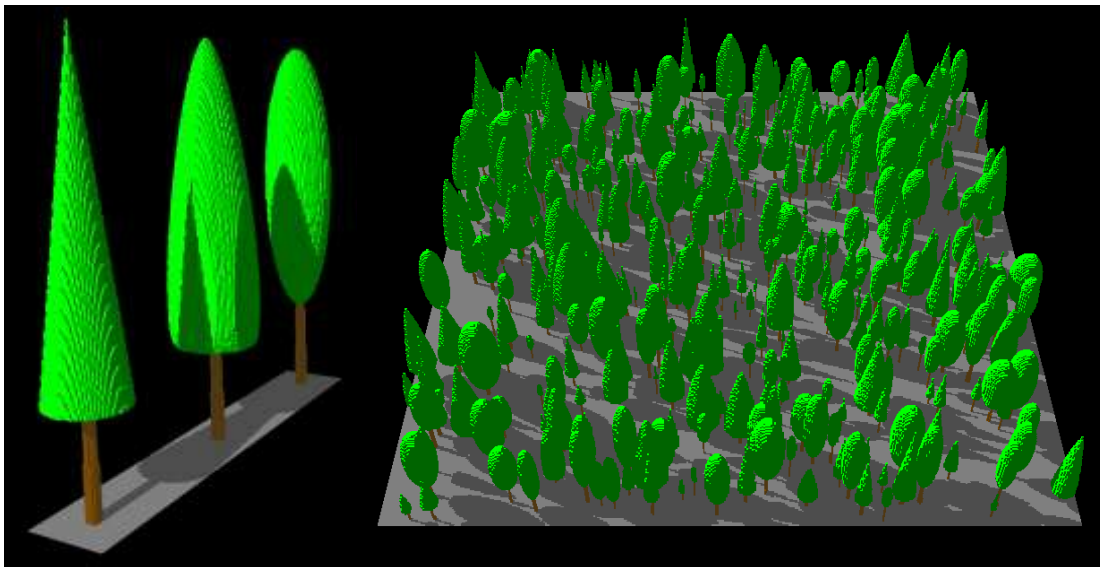
2.3 Modeled and LiDAR-based voxel scene representation

By using tree size information recorded during field visits of FIA plots, species-specific crown shape templates, and Forest Vegetation Simulator (Shaw 2009) formulae which estimate tree crown radius from stem diameter at breast height, 3D crown and stem models were generated for each tallied tree. The distribution of plot tree density was also assessed. Crown shape templates were designed as weighted combinations between cones and ellipsoids of predefined eccentricity. Stems were depicted as cones (Figure 2a). One hundred realizations of modeled tree stands, henceforth referred to as scenes, each 4 hectares in size, were then generated by sampling from the plot tree density distribution and the list of reconstructed crown shapes from the tallied trees (Figure 2b). The spatial distribution of stems in the modeled stands followed a Strauss (1975) process, a compromise between a regular point pattern (on a grid) and a complete random one. A Strauss point pattern is believed to approximate the spatial arrangement of trees in the Malheur National Forest. The geometric shapes of tree crowns and stems in the scene were then converted to a voxel representation. The latter is defined as the partitioning of space into discrete cubic elements. Voxels with center within a modeled stem or crown were labeled ‘filled’ otherwise they were labeled ‘empty’. Terrain in all modeled scenes was considered flat.

The 4-ha, LiDAR-based scene representations were centered on each of the 54 subplots. The buffer area around each subplot center ensures that the influence on illumination conditions of both the surrounding vegetation and the local topography is considered. Voxels containing one or more above-ground returns were labeled filled. Returns with above-ground elevation below 1 m were excluded from the voxel labeling process to ensure a clear distinction between ground and vegetation voxels. The former were identified using the DTM provided by the data vendor.

2.4 Voxel traversal by sun rays

Prior to actual voxel traversal operations, the illumination of cells in the subplot DTM is examined by application of the hillshade algorithm for all desired sun locations. Pronounced geomorphological features in the vicinity of the subplot such as mountain peaks, ridges, or steep



a. b.

Figure 2, a. Perspective view of modeled trees with, from front to back, conical, weighted and ellipsoidal crown shape; and b. Perspective view of modeled, 4-ha scene.

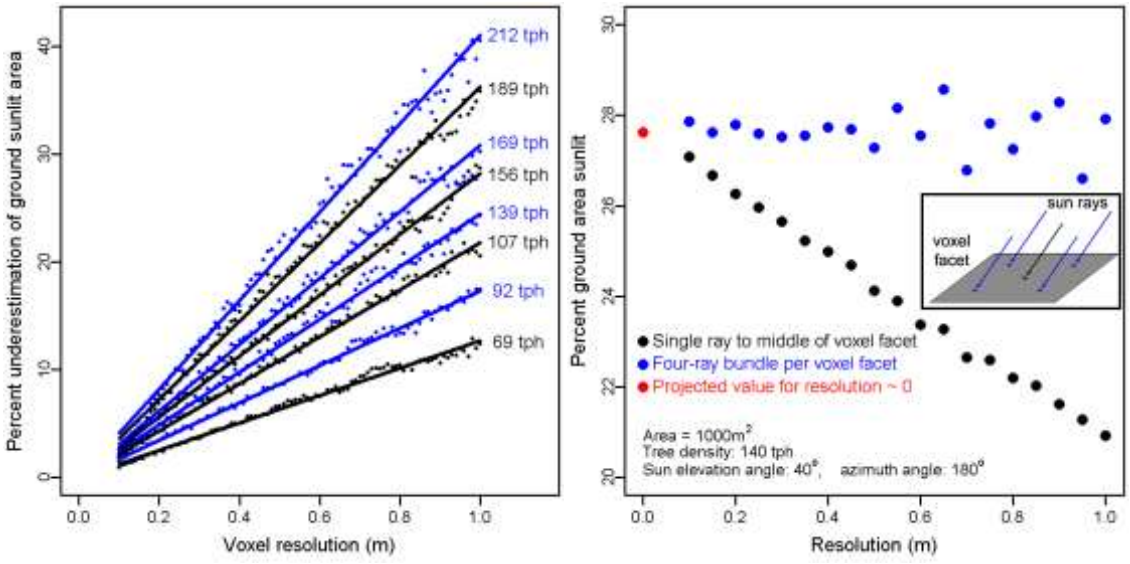
terrain could block incoming solar rays and if their presence is not considered sunlit conditions will be overestimated. After the absence of topography-induced shading effects has been confirmed, the directional vectors of the sun rays to the center of the parallelepiped enveloping the scene voxels are computed. The directional vectors are assumed constant within the scene considering its size in relation to the distance to the sun.

The voxel traversal by sun rays is a modification of the method by Amanatides and Woo (1987). For a given sun location, rays are generated towards the center of every visible voxel facet (side) in the scene. For each ray, the directional vectors calculated for the current sun location are used to determine the coordinates of the entry point to the scene. As the ray progresses towards the targeted voxel facet, its intersections with other voxel facets are computed and the label of the each voxel is examined. The ray terminates if it intersects the facet of a filled voxel. A facet is labeled sunlit only if it is reached by the ray directed to its center. The entire process, including the construction of the modeled scenes, was scripted in the C programming language with multithreading support. Sun azimuth and elevation angles for specific locations and times were computed by using a library compiled by Reda and Andreas (2008).

2.5. The role of voxel resolution

It was determined via experimentation with modeled scenes that, for a given solar position, as resolution becomes coarser, scene surfaces in shadow increase at the expense of those sunlit. The voxel resolution and sunlit or shadowed area relationship was found to be linear, and the underestimation of sunlit areas for a given resolution was more pronounced where tree density was higher (Figure 3a). This observation from the modeled scenes implies that unbiased assessment of light conditions via voxel representations, at least for crowns modeled as concave objects with continuous surfaces, is possible only with very fine voxel resolution. If applied to LiDAR-derived scenes, such a constraint would undermine the entire approach because, with too fine a resolution voxels, the representations of crowns would be exceedingly sparse even for high-density acquisitions. Hence, the sunlit areas would be grossly overestimated. Conversely, if resolution is too coarse, the volume occupied by crowns would be inflated and lead to overestimation of shadows. Thankfully, however, it was subsequently determined that consistently accurate estimates can be obtained even with coarser voxel resolutions. This was

accomplished by directing, instead of one, four sun rays towards each visible voxel facets in the arrangement seen in the inset of Figure 3b, and by labeling the facet sunlit as long as at least one ray were to reach it (Figure 3b). The four-ray per facet option was employed for all LiDAR-derived scenes.



a. b.

Figure 3, a. Relationship between voxel resolution and extent of sunlit areas in modeled stands as function of tree density, and b. effect of sun ray density and distribution [inset] on sunlit conditions.

Determining the resolution that yields accurate estimates of solar illumination regimes is therefore critical and dependent on the local density of the return cloud. Since return density seldom is stationary across the acquisition area, the optimal voxel resolution has to be determined dynamically. Following experimentation with independent field observation of light availability, the optimal voxel resolution was determined as the mean local spacing in tow dimensions among above-ground returns with elevation value within 2 m from the highest-return raster surface.

3. Results and Discussion

The resolution of the LiDAR data-derived voxel scenes centered on each of the 54 subplots with light observations ranged from 53 to 69 cm (mean 59 cm). The volume occupied by voxels containing at least one return from subplot tree crowns ranged between 25 and 65 percent of the volume that would have been occupied if the crowns of those trees had been modeled using the procedure described in section 2.3. Lower and internal crown components exhibited sparser representation.

Despite the voxel density and continuity issues for most crowns, a remarkable agreement between LiDAR-derived and field-measured sunlit illumination of ground surfaces was observed (Figure 4). The slope of field-measured values regressed linearly against the LiDAR-derived ones was not different from 1 at $\alpha = 0.01$. This finding indicates that the methodology employed is capable of delivering unbiased estimates of solar illumination regimes, at least in the presence of vegetation and LiDAR return density conditions similar to those evaluated in this study. Of the two outliers observed, the first with field index value approximately 0.5, or 25% direct sunlit illumination, belonged to an area with noted budworm infestation. In the year

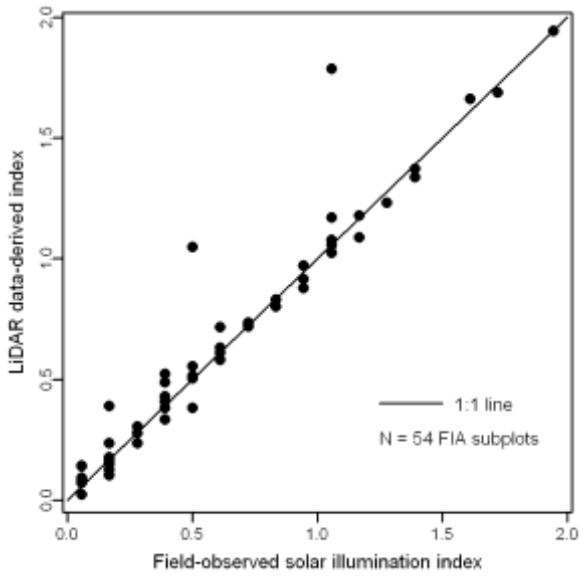


Figure 4, Comparison of LiDAR-derived estimates and field observations of ground solar illumination.

elapsed between LiDAR data acquisition and field light measurements, progression in infestation intensity and ensuing reductions in crown foliage could be responsible for overestimation of sunlit conditions. The second outlier was a subplot positioned at the edge of a forest stand towards a sizeable opening, where instability in aircraft attitude resulted in return density much lower than the mean across the entire acquisition. The ensuing undersampling of crowns at the edge of the stand likely led to overestimation of the light regime.

Visual qualitative assessment of LiDAR return clouds or even voxel representations suggests that ocular estimates of sun light illumination conditions can be substantially overestimated. In the scene shown in Figure 5 (a. and b.), only about 10 percent of the ground at the central half area of the scene is actually sunlit, an amount much less than the mean of 30 percent suggested by human cognition for the mid-morning of a summer day. Figure 5c illustrates the importance of using a buffer around the area of interest to negate scene edge effects. It was found that increasing the buffer width beyond 1.5 times the mean tree height did not improve the accuracy of sunlit area estimates. Cumulative estimates of sunlit conditions for a day (Figure 5d), or longer period, exhibited substantial variability and pronounced local spatial continuity.

The key concept of this study is the assumption that representations of tree crowns generated from LiDAR return clouds support analytical and explicit calculations regarding the timing and duration of solar illumination conditions within forest stands. With most recent LiDAR acquisitions in the State of Oregon obtained in high density and extending over large areas, the utility of dynamic light availability estimates evolves from a theoretical exercise to a practical management and decision making tool. The transition from concept to implementation required decisions on often elusive components that are integral to the implementation procedure such as the voxel resolution appropriate for each scene and number of rays to use. By using modeled representations of tree crowns, the effects of these components on the illumination estimates could be investigated, quantified, and ultimately controlled thereby ensuring estimation accuracy. Emphasis was given to multithreaded computational support which early on emerged as a logistical prerequisite, considering that each scene explored comprised over 7 million voxels, and the need to repeat the computations for series of solar positions.

The very good agreement between LiDAR-derived and field-recorded sun light conditions may not necessarily hold when the sun is low above the horizon. At low solar elevation angles, many

sun ray trajectories intersect exclusively the lower portions of crowns where the density of filled voxels is reduced. Since all field data were obtained within 4 hours from solar noon, the performance of this method could not be determined for early morning or pre-dusk times.

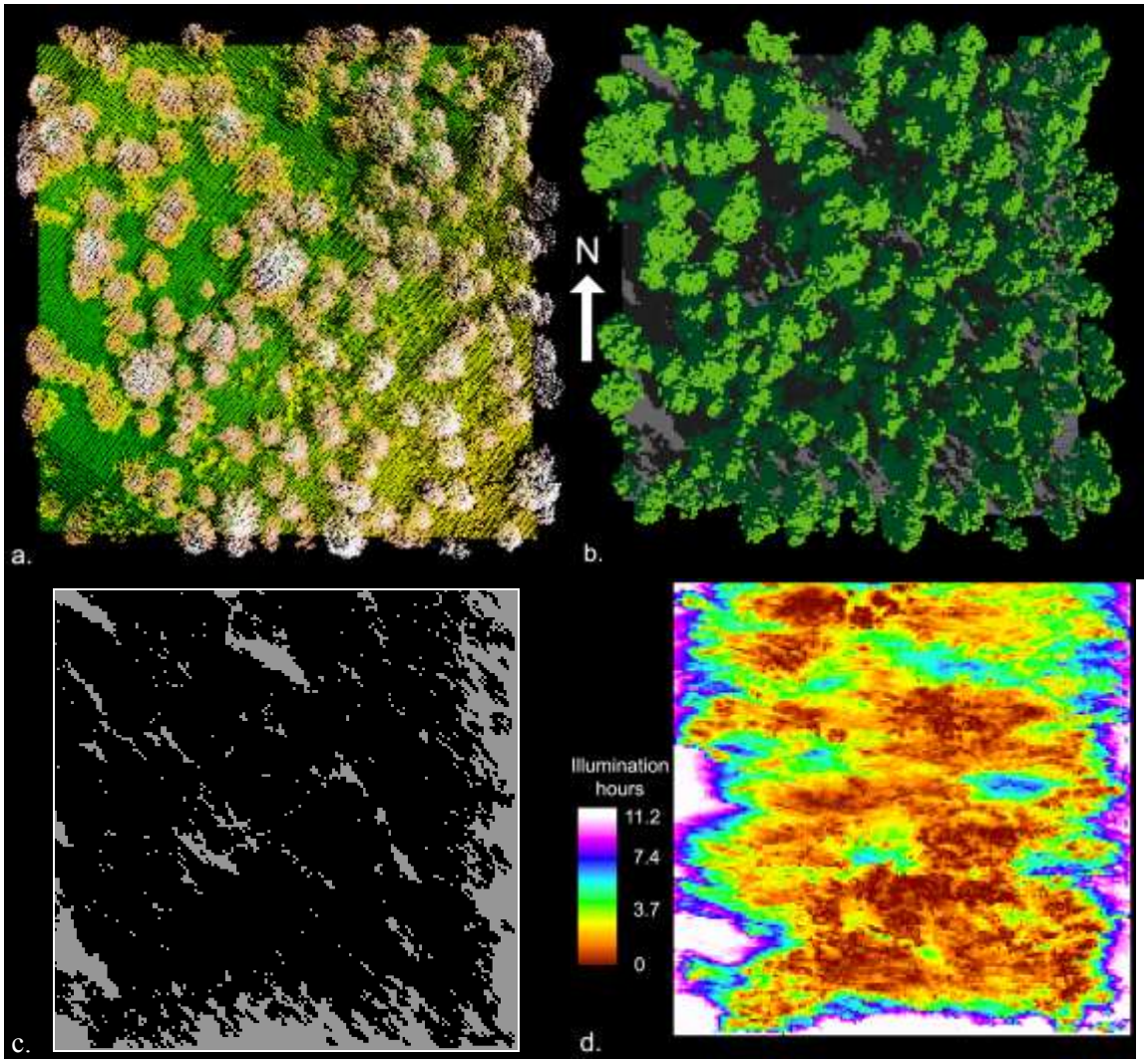


Figure 5, a. Nadir view of return point cloud, with light colors representing higher return elevation; b. corresponding nadir view of morning solar illumination conditions calculated and shown in voxel space; c. ground illumination conditions as in (b) but with vegetation not shown and light/dark tones representing sunlit/in shadow voxels respectively; and d. mean sunlit index of ground voxels for a mid-summer,12-hour period centered on solar noon.

At its present development stage, the traversal of LiDAR data-derived voxel scenes considers only direct sunlight, and tree crowns are assumed to be opaque objects. As it can be seen in Figure 2a, the shadow of modeled crowns cast on the ground or on adjacent crowns contains no sunlit portions. The suggested approach for dynamically determining the optimal voxel resolution of the LiDAR-derived scenes also aims at generating opaque crowns without any, or with minimal openings within them. Since light transmittance through the canopies is not yet considered, all sunlit indices calculated lack the diffuse solar illumination component.
This limitation, however, can likely be addressed by directing towards each visible facet a predefined distribution of regularly spaced rays in a hemispherical arrangement. The ratio of rays that reach the facet to the total would quantify scene openness around the facet and serve as surrogate of diffuse lighting conditions. Considering that scene openness is not subject to

diurnal cycles, the computation load for obtaining both direct and diffuse sunlight illumination conditions could be manageable. The diffuse illumination component could then be added to the indices.

Obtaining an adequate number of precisely timed and georeferenced sunlight availability field observations in various stand density and morphology conditions is a prerequisite for evaluating and enhancing the methodologies described but also a major challenge. In addition to logistical cost, illumination conditions often change too rapidly for a human observing a selected number of points. Small ground sunspots generated by sun rays propagating through crowns flicker substantially even with the lightest wind. The point observations used in this study are better suited to quantifying direct sunlight illumination where there is a distinct boundary between sunlit and shadowed areas. Relying on quantum sensors for an instantaneous measurement would likely underestimate direct illumination since the instruments are designed for operation during overcast conditions only. Alternative data collections strategies would hence need to be devised for measuring direct and diffuse illumination conditions simultaneously.

4. Conclusions

This study uses LiDAR data-derived representations of tree crowns and terrain in voxel space and ray tracing to estimate direct sunlight illumination. The effect of voxel resolution is investigated using simulated scenes, also in voxel space. Suggestions on how voxel resolution conducive to unbiased estimation of illumination conditions can be calculated are offered. Remarkable agreement between field measurements of sunlight availability and corresponding LiDAR-data based estimates was observed. Distributed estimates of light transmittance thought canopy or light availability in diffuse form, although not yet implemented in voxel space, are likely attainable and will be investigated in future research. Since the investigation was conducted in low to medium canopy cover conditions and gentle terrain, additional experimentation would be required to determine the approach's performance in close-canopy forests or in the presence of steep slopes.

References

Amanatides, J., and Woo, A. 1987. A fast voxel traversal algorithm for ray tracing. In G. Marechal (Ed.). *Proceeding of EUROGRAPH-ICS'87*, North-Holland, Elsevier Science Publishers: 3-9.

Bellow, J.G., and Nair, P.K.R. 2003. Comparing common methods for assessing understory light availability in shaded-perennial agroforestry systems. *Agricultural and Forest Meteorology*, 114, 197-211.

Burrough, P.A. and McDonell, R.A., 1998. *Principles of Geographical Information Systems*, Oxford University Press, New York, p. 190.

Brunner, A. 1998. A light model for spatially explicit forest stand models. *Forest Ecology and Management*, 107, 19-46.

Chen, Y.D., Carsel, R.F., McCutcheon, S.C., and Nutter, W.L. 1998. Stream temperature simulation of forested riparian areas: I. Watershed-scale model development. *Journal of Environmental Engineering*, 124(4), 304-315.

Comeau, P.G., Braumandl, T.F., and Xie, C.Y. 1993. Effects of overtopping vegetation on light availability and growth of Englemann spruce (*Picea engelmannii*) seedlings. *Canadian Journal of Forest Research*, 23, 2044-2048.

MacFarlane, D.W., Green, E.J., Brunner, A., and Amateis, R.L. 2003.Modeling loblolly pine canopy dynamics for a light capture model, *Forest Ecology and Management*, 173(1–3), 145–168.

Nicotra, A.B., Chazdon, R.L., and Iriarte, S.V.B. 1999. Spatial heterogeneity of light and woody seedling regeneration in tropical we t forests. *Ecology*, 80(6), 1908-1926.

Parendes, L.A., and Jones, J.A. 2000. Role of light availability and dispersal in exotic plant invastion along roads and streasm in the H. J. Andrews Experimental Forest, Oregon. *Conservation Ecology*, 14(1), 64-75.

Parent, S., and Messier, C. 1996. A simple and efficient method to estimate microsite light availability under a forest canopy, *Canadian Journal of Forest Research*, 26(1), 151–154.

Shaw, J.D. 2009. Using FIA data in the Forest Vegetation Simulator. In: W. McWilliams, G. Moisen, R. Czaplewski (Eds). *Forest Inventory and Analysis (FIA) Symposium* 2008; October 21-23, 2008; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 16 p.

Strauss, D.J. 1975. A model for clustering. *Biometrika*, 63, 467–475.

Reda, I., and Andreas, A. 2008. Solar Position Algorithm for Solar Radiation Applications. Technical Report NREL/TP-560-34302, National Renewable Energy Laboratory, Golden, Colorado, 56 pp.

Cartography of uncertainty of predicted forest variables estimated from ALS data

Francisco Mauro[1], Rubén Valbuena[2] David García García[3] & María Victoria Nuñez[4]

[1]Francisco Mauro Gutiérrez. Scholarship researcher. ETSI Montes. Polytechnic University of Madrid.Ciudad Universitaria SN. 28040 Madrid francisco.mauro@upm.es
[2]Rubén Valbuena Puebla. Scholarship Researcher. European Forest Institute HQ Torikatu 34 80100 Joensuu, Finland ruben.valbuena@efi.int
[3] David García García. Scholarship researcher. ETSI Montes. Polytechnic University of Madrid.Ciudad Universitaria SN. 28040 Madrid dgarcia.itf@gmail.com
[4]María Victoria Martí Núñez. Research Assistant. ETSI Montes. Polytechnic University of Madrid.Ciudad Universitaria SN. 28040 Madrid mvn@iies.es

Paper Number: SL2012-178

Abstract

Linear regression is probably the most extended technique to predict forest attributes from Airborne Laser Scanner (ALS) data. It frequently happens that the model that minimizes the residual variance (Type 1 model) violates assumptions about the residuals that allow controlling the conditional uncertainty. In such situations arise a question about whether or not it is preferable to discard the Type 1 model and search for a new model which verifies the entire above mentioned hypothesis but with a higher residual variance (Type 2 models). By controlling the conditional variance it is possible to report uncertainty measures at pixel level as accuracy maps which can be of interest if decisions are involved at pixel scale. In this study we analyzed the differences of obtaining either regression models in which global accuracy is optimized or obtaining models in which uncertainty is fully controlled when predicting basal area (G), stand density (N), dominant height (Ho), mean tree height (Hm), standing volume (V), tree biomass (B) and quadratic mean diameter (Dg) from ALS data. Models of Type 1 and 2 were fitted for the above mentioned variables and their logarithms. The fitting was repeated 250 times for each variable using samples of 37 plots taken with replacement form the original set. The RMSE for models of Type 1 for G, N, Ho, Hm, V, and B were 2.08%, 9.97%, 2.97%, 11.43%, 3.80% and 5.20% lower than the RMSE for models controlling the conditional variance, Type 2.

1. Introduction

At the present, forest attributes are predicted from Airborne Laser Scanner (ALS) in operational applications that go beyond academic or research environments. Most of the applications of this technology to forest inventory have followed regression based approaches in which forest attributes are predicted from ALS-derived metrics. This method is usually referred as “Area Based Approach” (ABA), and it focuses on finding empirical relationships between ALS derived predictors and forest properties (Erdody and Moskal 2010, Maltamo et al. April 2006, Næsset 2002). Another type of approach sometimes called, “Individual Tree Approach”, has also been widely used with high density ALS data. This approach is based on finding trees in the ALS data and using these data for estimating the more interesting properties of the located trees (Hirata et al. 2009). Finally, a third way of deriving forest attributes from ALS data, that focuses on modeling the physical interactions between the forest canopies and the laser pulses has also been employed (Model Based Approach) (Magnussen et al. 1999, Mehtatalo and Nyblom 2012, Mehtatalo and Nyblom 2009).

The area based approach in combination with linear regression analysis is probably the most extended technique. Most of the studies, relating forest attributes and ALS data using this technique, minimize and report estimates of the marginal (i.e. global), uncertainty of the predictive models (Type 1 models). It frequently happens that the model that minimizes the residual variance violates assumptions about the residuals that allow controlling the conditional (i.e. pixel level) uncertainty. These assumptions are:

1. Linear relationship between predictors and response
2. The residuals are independent
3. The residuals are i.i.d and follow a Normal distribution with mean 0 and constant variance.

If the hypotheses are not met, marginal uncertainty provides little information about conditional uncertainty. This fact makes it impossible to estimate the uncertainty for specific predictions at pixel level. In scenarios of small scale forestry, where the pixel size can constitute a treatment, logging, or management unit, predictions at pixel level might be one of the inputs that the decision maker would use to orient the planning. In addition, any potential user of the predictive maps would expect differences between the predictions and subsequent field measurements being in the range of the reported uncertainty. If the assumptions are wrong, strong differences with the reported uncertainty can appear and make the final users lose confidence in the provided map. This problem can be accentuated by the fact that in most cases differences between model and measurements are larger for the most valuable products. For example, when predicting standing volume (V), uncertainty is usually small for areas with low quantities of wood, which are of low merchant value, while these residuals are, in general, large for areas with large standing volumes and value. Therefore, providing only global estimates of the uncertainty of the predictive models can be misleading. An appealing aid for decision makers or final users would be to have pixel level measures of the uncertainty of the predictions.

Several techniques, such as transformations of the response variables or Generalized Linear Models, have been used to deal with the heterogeneity of the residuals and/or to meet other assumptions. Making transformations of the response variable is the most extended technique. logarithmic, power function, Box-Cox and root square transformations have been used in several studies (Næsset 1997, Næsset 2002, Andersen et al. 2005, González-Ferreiro et al. 2012). By changing the scale of the response variable, sometimes, it is possible to find Type 2 models. In such cases, the uncertainty would be controlled in the transformed scale. However, for certain transformations it is also easy to control the uncertainty when transforming back to the original scale. That is the case for the logarithmic transformation.

When the assumptions are met it is possible to obtain confidence intervals for predictions, which are conditional estimates of the uncertainty at pixel level. These uncertainties can be reported for each pixel as lower and upper limits of confidence intervals. The result is a pair of ancillary accuracy maps which can be a valuable tool for decision makers who will use the prediction map. These maps are easier to understand for final users and provide a good image of the possible differences that can be found between map and further field measurements. However, pixel level estimates of the uncertainty require obtaining predictive models in which residuals must verify the abovementioned hypotheses. Verifying these hypotheses may lead to models (Type 2 models) in which the marginal or global uncertainty is not optimized.

In order to decide what type of model is better suited for a specific purpose, questions about the tradeoffs of, using models in which global accuracy is optimized or models in which uncertainty is fully controlled, need to be answered. The aim of this work is to evaluate the losses in global accuracy when obtaining linear models of Type 2, that permit computing accuracy maps, instead of predictive models in which marginal uncertainty is optimized without controlling the

conditional variance (Type 1).

2. Method

2.1 Study area

The study area is a 300-ha Pinus sylvestris forest located in Guadarrama Mountains (Central Spain). A total of 37 circular plots, of 20 m radius, were systematically distributed on three lines starting at randomly selected locations close to points with easy access. Plots were georreferenced using a phase differential GPS technique. GPS/GLONAS observations, of least 30 minutes, were taken at each plot center with a logging rate of one observation per second. diameter at breast height (DBH) and tree height were measured for every tree in the plot. Basal area (G), stand density (N), dominant Height (Ho), mean tree height (Hm), and quadratic mean diameter (Dg) were directly computed from the field measurements. Standing volume (V) and tree biomass (B) were computed for each tree using local allometric models and then aggregated at plot level. The model for V considered the DHB and the tree height as explanatory variables (Rojo and Montero 1996) while the model for B only considered the DBH (Mc Gaughey 2011).

2.4 ALS data

ALS data was acquired in 2006 using an ALS50-II scanner from Leica Geosystems (Switzerland). The Field of View was 25º and the average flying altitude was approximately 1500 m above the ground level. The pulse rate was 55 kHz, and a maximum of 4 returns per pulse were recorded. For each return, intensity measurements were recorded. The final point density was higher than two points/m^2. Irregularities in point density were removed obtaining a final and homogeneous density of two points/m^2. A detailed description of the ALS data can be found in (Valbuena et al. 2012).

The raw point cloud was classified using Terrascan (Terrasolid, Finlad). Points classified as ground points were used to generate a 1 m resolution Digital Terrain Model (DTM) using krigging interpolation. This DTM was employed to normalize the point cloud. Height and Intensity metrics, (percentiles, maximums, minimums, mean values etc..) and density metrics, (the number of first, second, third and fourth returns…) were computed for each plot using FUSION (Mc Gaughey 2011). These variables and their logarithms, a total of 124 variables, were considered as potential predictors.

2.3 Model fitting

The number of potential predictors was very high; for this reason, a method for selecting simple enough models was implemented in R (Team 2011). For a response variable the procedure started selecting the 15 strongest correlated predictors. When correlation of any of these 15 variables and the response was not significant the predictor was discarded. Then these new set of 15 or less predictors was used in a second stage. This stage used the method *leaps()* from the R package *leaps*(Lumley 2009) . This method takes as input a response Y variable and a set of predictors X. It performs an exhaustive search for the best subsets of the variables in X for predicting the response Y using linear regression. The output in this case was a list in which, the best 5 combinations of variables was reported for every number of predictors from 1 to the number of variables in X. Due to the reduced sample size, sets of more than 4 variables were discarded. Collinearity among predictors was tested in the remaining set of combinations of variables, and those in which the condition number κ were larger than 30 were also removed. The groups of variables were sorted according to their adjusted r-squared (adj R^2) to penalize the complexity of the models. The result of this procedure was a sorted list with up to 20

combinations of variables. The first group of variables in the list was considered the best for reducing the marginal uncertainty and developing models of Type 1. Finally, for models of Type 2 an additional step was performed. We searched in the previous list for the model with the highest adjusted-r squared that also verified that residuals:

- Were normally distributed.
- Showed a constant variance
- Their mean could be considered 0.

These three hypotheses were respectively tested using the Shapiro-Wilks test, the Breusch-Pagan Test and the student-t test. The independence of the residuals was accepted beforehand as the distance between plots was considered large enough to avoid spatial correlation. The first group of variables that also verified these hypotheses was considered the best for obtaining models of Type 2. If no model of Type 2 was feasible, the method returned a special value for signaling. Due to the reduced number of plots, the root-mean-squared error RMSE for each model, either of Type 1 or of Type 2, was computed by leave one out cross validation. Logarithmic transformations were taken into account because the distribution of the residuals was easy to control when transforming back to the original scale. Those magnitudes following a Normal distribution in the logarithmic scale would follow a log-Normal distribution when transformed back. When considering log-transformed variables, the bias for back-transformed predictions was corrected as in (Montero et al. 2006) and the RMSE was computed in the original scale of the variable.

Models of Type 1 and Type 2 were fitted for the above mentioned forest variables and their logarithms using the explained method. The result was in a group of four models. The fitting was repeated 250 times for each variable using samples of 37 plots taken with replacement from the original set of plots. The best model of Type 1 and the best Model of Type 2 when considering both, the original responses and their transformations, were compared. The relative loose in RMSE (*RL_RMSE*) when obtaining Type 2 models was computed using equation 1 for each iteration:

$$RL\_RMSE = 100(1 - RMSE_{b2}/RMSE_{b1}) \quad (1)$$

Where $RMSE_{b1}$ and $RMSE_{b2}$ are the RMSE of the best model of Type 1 and Type 2 respectively.

The number of times in which it was impossible to obtain models of Type 1 or models of Type 2 was recorded and the proportion to the total of iterations was computed.

Models of Type 2 allow computing confidence interval maps. If no transformation is used, pixel levels estimates of these intervals can be computed using equation 2. If a logarithmic transformation is used the bounds of the intervals should be transformed back to the original scale.

$$y_x - \hat{y}_x | X_0 \sim t(n - p - 1, \sqrt{\sigma^2 - \sigma^2 X_0 X X^{-1} X_0'^2}) \quad (2)$$

Where $y_x - \hat{y}_x$ is the difference between model and prediction and a possible field value, n is the number of observations, p is the number of variables in the model, X_0 the values of the predictors in a specific point, σ^2 is the residual variance and should be replaced by an estimate of it, and X is the design matrix.

3. Results

RMSE for models of Type 1 for G, N, Ho, Hm, V, and B were 2.08%, 9.97%, 2.97%, 11.43%, 3.80% and 5.20% lower than the RMSE for models controlling the conditional variance, Type 2 (Table 1). The percentage of iterations in which it was impossible to fit models of Type 1 or Type 2 is also shown in Table 1.

Figures 1(a,b,c) show a Type 2 prediction map for the standing volume, the corresponding upper and lower limit for the confidence interval of the prediction at pixel level. Half of the width of the intervals, relative to the prediction, is shown in Figure 1 (d). The model showed an R^2 of 86.54 % and the RMSE was 72.13 m^3/ha.

Figure 1: (A) Prediction map from Type 2 for V. (B,C) Lower and upper limit for the confidence interval of the predictions. (D) Half width of the confidence interval relative to the prediction value.

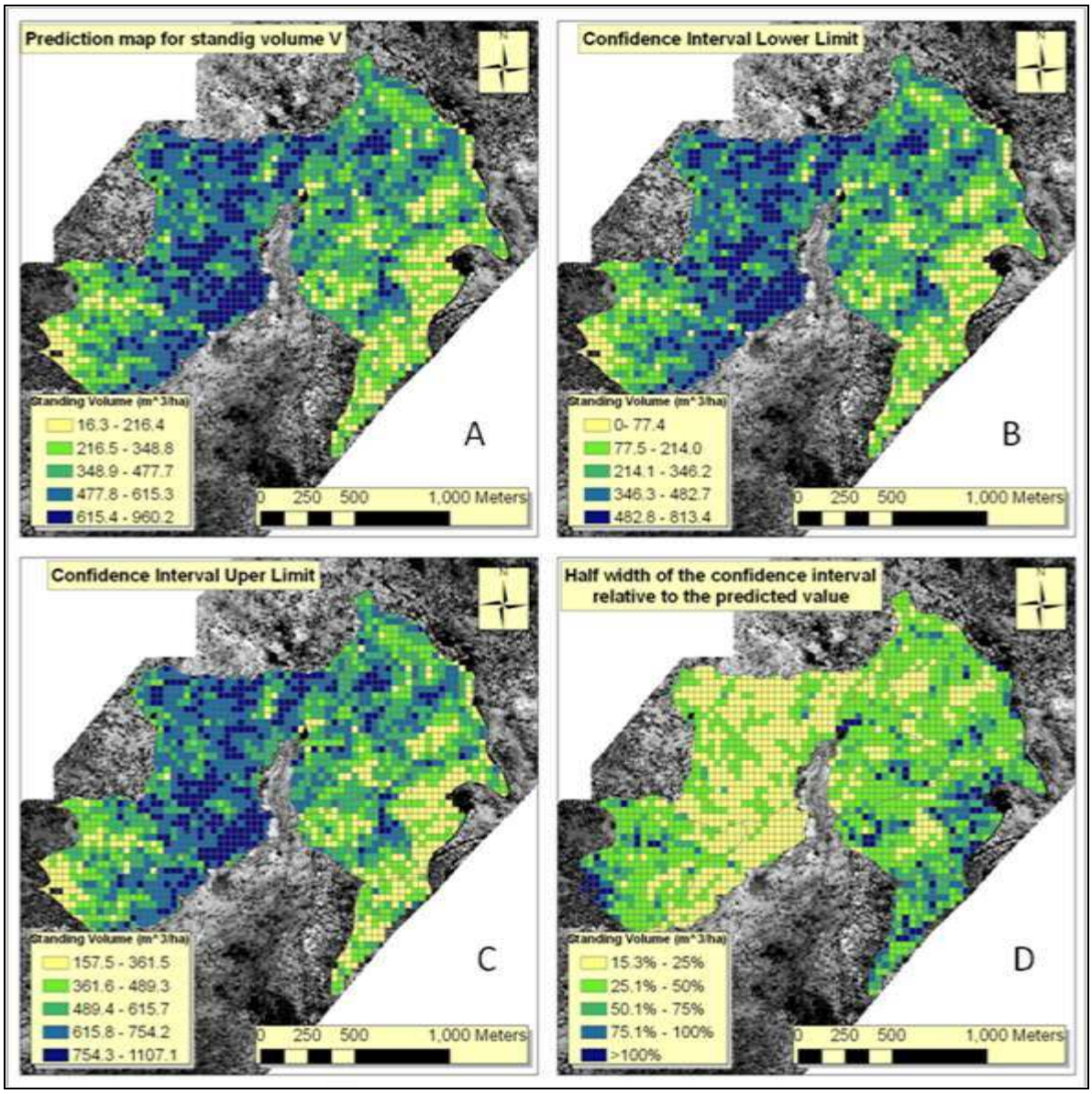


Table 1: Looses in RMSE when fitting models of Type 2 and percentage of fails when fitting models of Type 1 and Type 2.

Variable	RL_RMSE	% Fails obtaining Type 1 models	% Fails obtaining Type 1 models
G	2.08%	0.00%	0.00%
N	9.97%	26.40%	31.20%
Ho	2.97%	22.80%	47.60%
Hm	11.43%	0.00%	71.20%
V	3.80%	0.00%	0.00%
B	5.20%	0.00%	0.00%

4. Discussion

When minimizing global accuracy RMSE is reduced in less than an 11.5% for any of the considered variables and this reduction is smaller than a 5.5% for the majority of them. Developing models Type 2 may be of interest as of obtaining accuracy maps can compensate for the small reduction in global accuracy. Fitting models of Type 2 can be of little interest when the objective is providing large scale estimates of forest variables. However, forest managers or decision makers may require estimates of variables of interest in small areas. For example, the average size of the private forest property in Spain is very small, being this medium size less 10 ha in Cataluña, Galicia and the Basque Country and estimates are needed for those small properties. In such conditions, decisions to be supported by estimates for areas as small as the plots considered in this study would benefit from a deeper knowledge of the uncertainty of the predictions.

García et al. (2010) developed models of Type 2 for estimating G, N, V and B in Cuenca Spain. As a product from the model confidence interval were computed for each pixel. The bounds of these intervals were reported as accuracy maps similar to those in Figure 1b, and 1c. In this case the election models of Type 2 and the elaboration of accuracy maps was motivated by the aim of producing understandable maps in which the uncertainty linked to each pixel of the prediction maps was clearly illustrated.

The reported decrease in RMSE can be considered as a reference for future applications in similar environments. For example, in a hypothetical case of developing regression models for predicting the standing volume in close area, after obtaining models of Type 1, one can guess if obtaining model of Type 2 would provoke an important decrease in accuracy. The expected increase in RMSE would be a 3.8%, which seems to be reasonably small.

The model used as an example for Figure 1 had an RMSE and an R^2 similar to those in the model found by González-Ferreiro et al. 2012 for *Pinus radiata*. Even though the RMSE and the R^2 are relatively high, differences between prediction and observed values observed values are likely to be larger than a 50% of the predicted values. (Figure 1d). If this map was used for planning forest operation at a large scale, these errors would be averaged and the impact of the uncertainty at pixel level would be small. However, if the planning scale is small (areas of a small group of pixels, for example 2ha), the forest manager should take in account the uncertainty maps as the differences between field values an predictions can be high.

5. Conclusion

Accuracy maps are an understandable way to communicate the uncertainty of predictive

models and they facilitate a correct interpretation of predictive maps by final users. Reporting the accuracy of a predictive map can be a very valuable for small scale forest planning as the information provided by the accuracy maps can compensate for the decrease in global accuracy.

Acknowledgements

This work has been supported by Polytechnic University of Madrid.

References

Andersen, H., McGaughey, R.J., and Reutebuch, S.E. 2005. Estimating forest canopy fuel parameters using LIDAR data. Remote Sens. Environ. **94:** (4) 441-449. xxx10.1016/j.rse.2004.10.013

Erdody, T.L., and Moskal, L.M. 2010. Fusion of LiDAR and imagery for estimating forest canopy fuels. Remote Sens. Environ. **114:** (4) 725-737. xxxDOI: 10.1016/j.rse.2009.11.002

García, D., Godino, M., and Mauro, F. 2010. Estimación de variables de interés forestal basada en datos Lidar en el monte número 117 del C.U.P. término municipal de Cuenca. Available from http://oa.upm.es/6171/2/PFC_DAVID_GARCIA_GARCIA_TEXTO.pdf.

González-Ferreiro, E., Diéguez-Aranda, U., and Miranda, D. 2012. Estimation of stand variables in Pinus radiata D. Don plantations using different LiDAR pulse densities. Forestry xxx10.1093/forestry/cps002

Hirata, Y., Furuya, N., Suzuki, M., and Yamamoto, H. 2009. Airborne laser scanning in forest management: Individual tree identification and laser pulse penetration in a stand with different levels of thinning. For. Ecol. Manage. **258:** (5) 752-760. xxx10.1016/j.foreco.2009.05.017

Magnussen, S., Eggermont, P., and LaRiccia, V.N. 1999. Recovering Tree Heights from Airborne Laser Scanner Data. For. Sci. **45:** 407-422(16).Available from http://www.ingentaconnect.com/content/saf/fs/1999/00000045/00000003/art00012

Maltamo, M., Eerikäinen, K., Packalën, P., and Hyyppä J. April 2006. Estimation of stem volume using laser scanning-based canopy height metrics. Forestry} **79:** (2) 217-229. xxx10.1093/forestry/cpl007 Available from http://forestry.oxfordjournals.org/content/79/2/217.abstract

Mc Gaughey, R.J. 2011. FUSION/LDV: Software for LIDAR Data Analysis and Visualization.

Mehtatalo, L., and Nyblom, J. 2012. A Model-Based Approach for Airborne Laser Scanning Inventory: Application for Square Grid Spatial Pattern. For. Sci. **58:** (2) 106-118. xxxdoi:10.5849/forsci.10-023" Available from http://www.ingentaconnect.com/content/saf/fs/2012/00000058/00000002/art00002. [accesed 20 June 2010].

Mehtatalo, L., and Nyblom, J. 2009. Estimating Forest Attributes Using Observations of Canopy Height: A Model-Based Approach. For. Sci. **55:** (5) 411-422.Available from http://www.ingentaconnect.com/content/saf/fs/2009/00000055/00000005/art00004. [accesed 20 June 2010].

Montero, G., Ruiz-Peinado, R., and Muñoz, M. 2006. Producción de biomasa y fijación de CO2 por los bosques españoles. 1st ed. INIA, Madrid.

Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. **80:** (1) 88-99. xxxDOI: 10.1016/S0034-4257(01)00290-5

Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing **52:** (2) 49-56. xxxDOI: 10.1016/S0924-2716(97)83000-6

Rojo, A., and Montero G. 1996El pino silvestre en la Sierra de Guadarrama [In Spanish].

inisterio de Agricultura Pesca y Alimentación, Madrid. . 221

Team, R.D.C. 2011. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Lumley, T. 2009. leaps: regression subset selection. Available from http://CRAN.R-project.org/package=leaps.

Valbuena, R., Packalén, P., Martı´n-Fernández, S., and Maltamo, M. 2012. Diversity and equitability ordering profiles applied to study forest structure. For. Ecol. Manage. **276:** (0) 185-195. xxx10.1016/j.foreco.2012.03.036

Mapping and Managing the Fuzzy Forest

Günther Bronner
Umweltdata GmbH Mödling, Austria; g.bronner@umweltdata.at

Paper ID: SL2012-193

1. Introduction

Forest Management Plans (FMP) are composed from tables, maps and reports. Mapping the forest usually means to segment the forest area into vector polygons in order to get corresponding units in both: maps and tables. A segmented unit ('subdepartment') should be homogeneous in description, it should be appropriate for management as well. In Austria, the tradition of drawing lines on forest land to separate different stands goes back to the roots of Forest Management Planning in the 18th century.

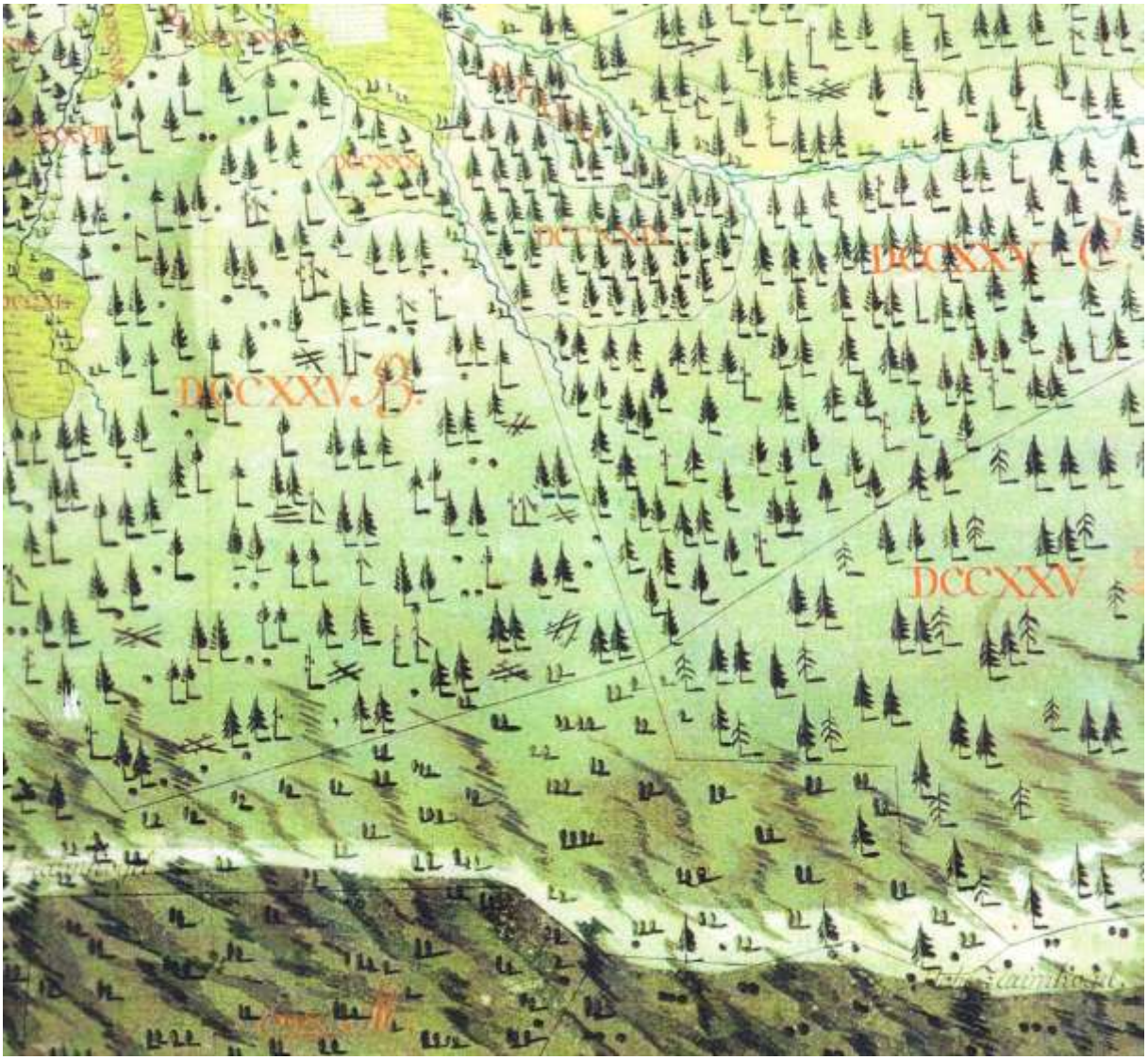


Figure 1: Detail from a forest map from 1759, Grundlsee, Austria (colored ink drawing)

Segmenting the forest land into subdepartments came along with a clear-cut harvesting strategy: After clear-cutting an area, followed by reforestation, the area was definitely homogeneous regarding age and tree species. An even-aged forest is easy to describe. Since decades, the preferred type of forest maps in Austria shows 20-year-age-classes. This kind of generalization was welcome, but it comes to its limits.

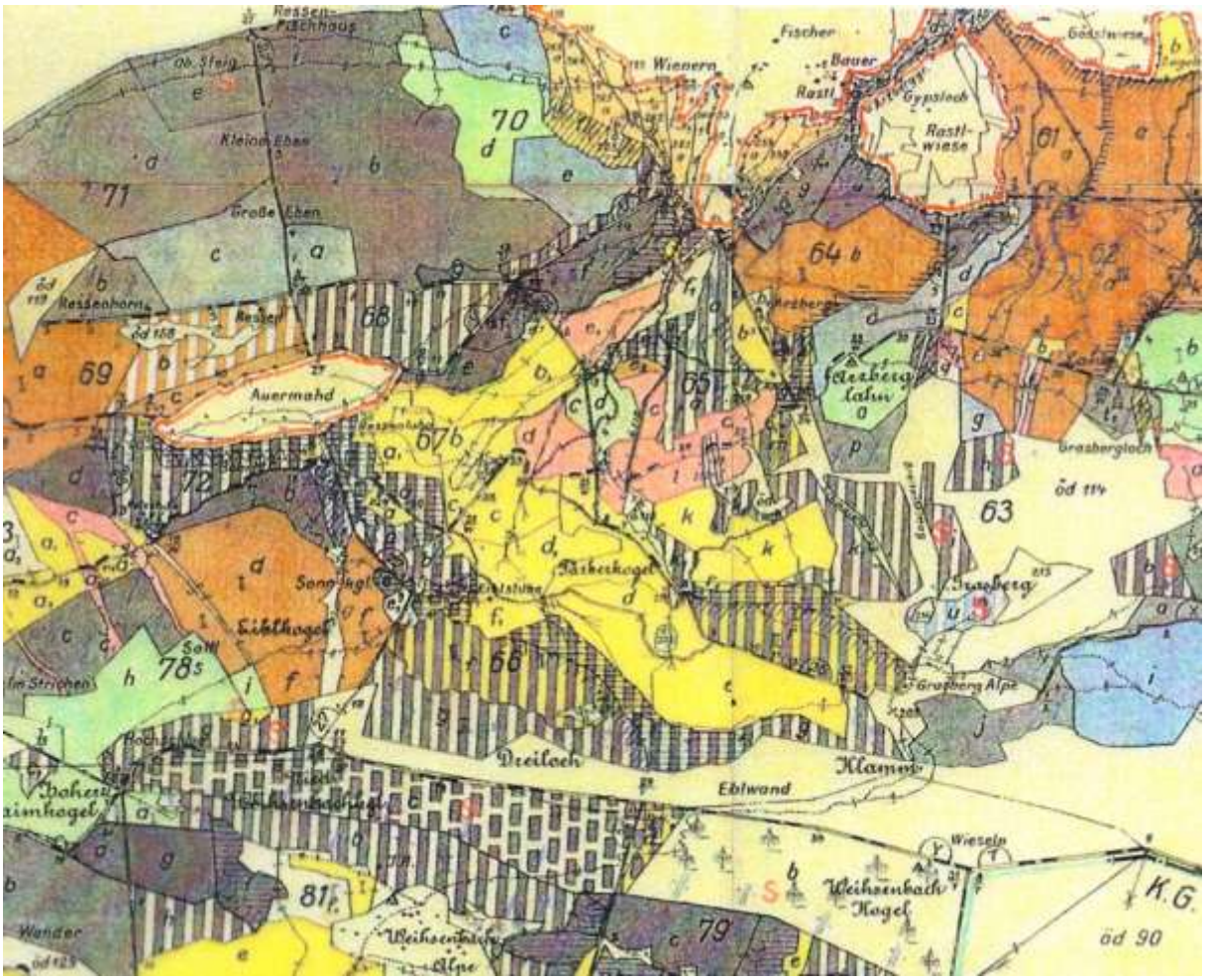


Figure 2: Detail from a traditional forest map, 1950, Grundlsee, Austria (black print, ink colored)

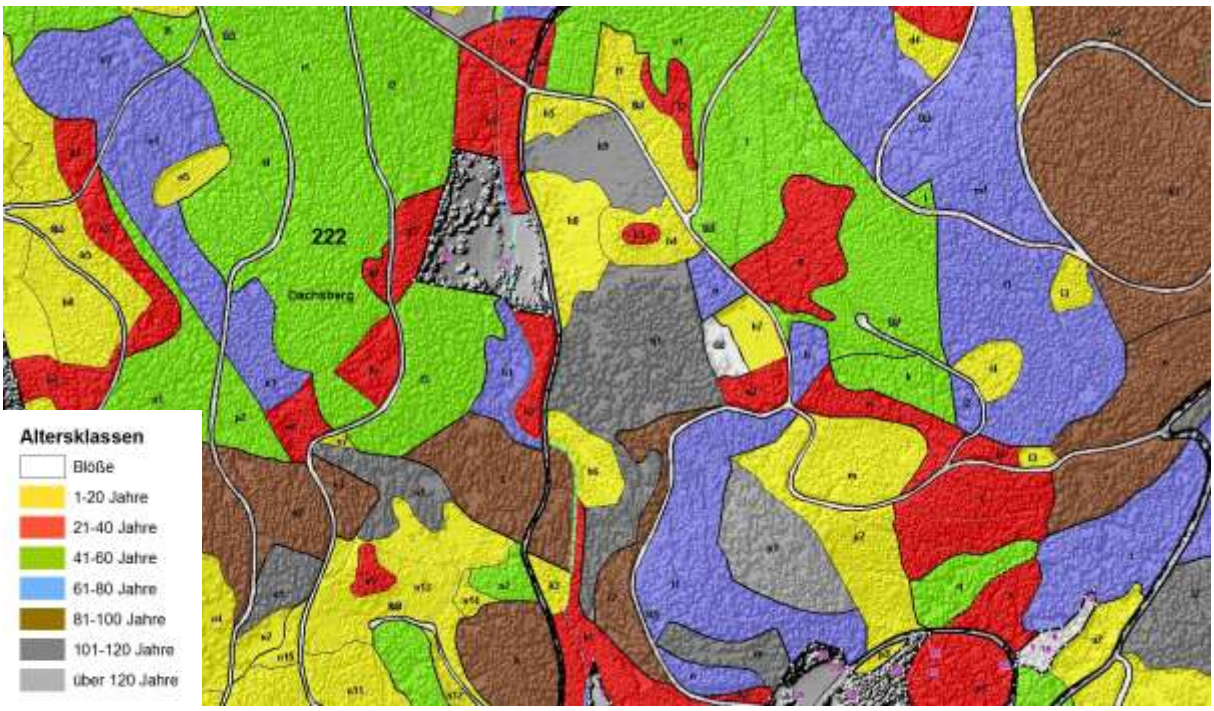


Figure 3: Detail from an actual GIS-designed map, showing age-classes and a canopy shaded relief

Within all properties of forest areas, age class and tree species are the only ones that could be indicated as discrete, while relief, geology, tree heights, growth rates, density, stock volume, they all change continuously in most cases. After some decades of nature-near selection-felling methods, even age and species of trees have become attributes with smooth transition. **FuzzyForest** will be my term for those forest areas, which resist any attempt of a reasonable segmentation.

In Figure 4 and 5 the difference can be realized: even-aged forest with segmentation into subdepartments and uneven-aged forest, where segmentation does not any more make sense. Both maps show a shaded relief of the canopy with colors from the canopy elevation and the lines for segmentation. In figure 5 the former segmentation lines are quiet arbitrary: most of them can not be identified in the forest.

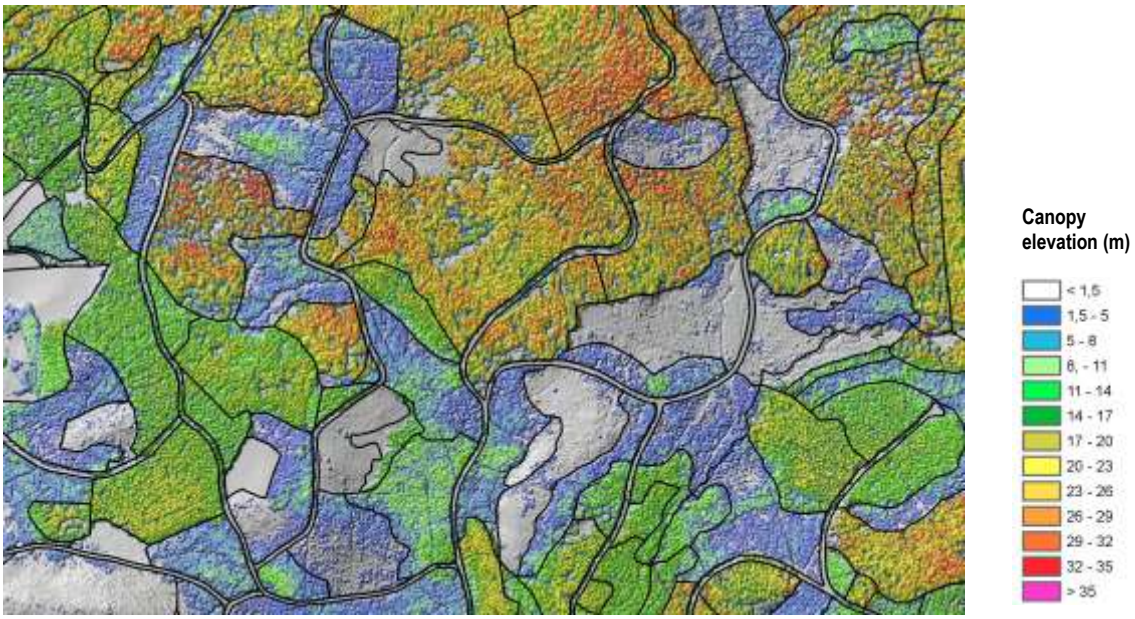


Figure 4: fusion of shaded relief with colors from canopy elevation in an even-aged forest of norway spruce.

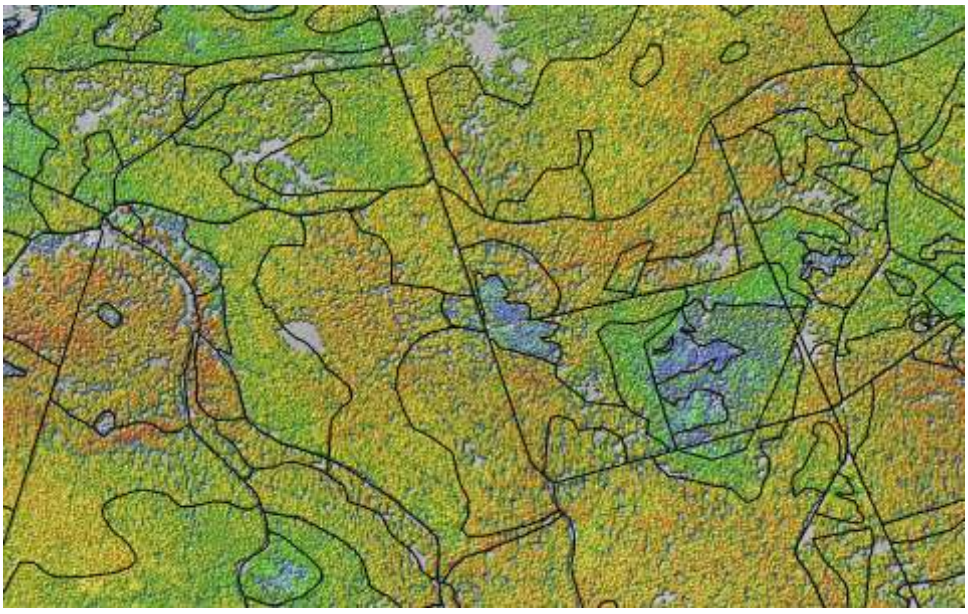


Figure 5: fusion of shaded relief with colors from canopy elevation in an uneven-aged forest of norway spruce.

Data from Airborne Laser Scanning (ALS) allow new methods of mapping, and they will change the traditional ways of management. In order to develop mapping methods and managing techniques applicable to FuzzyForests, the starting position will be summarized like this:

- A forest ranger, running the day-to-day business in forestry, needs significant maps showing manageable units. In Austria those units are usually 0.5 to 20 hectares large.
- Up to now, segmentation of forest areas into homogeneous subdepartments provided such manageable units.
- In uneven-aged forest areas, segmentation will not produce suitable results; the application of high sophisticated segmentation tools on ALS-data (like econition ®) will not change this insight. The resulting segmentation lines can often not be identified in the nature.
- Forest rangers prefer / are used to generalized maps and some basic figures of planning (harvesting) units, like an estimated stock volume, growth-rates, age and species of trees.
- Grid-based inventory methods provide those figures for larger units with high accuracy.
- ALS-based methods may provide figures for smaller local units with adequate accuracy.
- ALS-data allow new mapping methods that could meet the requirements of forest rangers (and thereby replace the beloved age-class-maps)
- Tablet PC's could provide multiple maps as well as planning figures from preprocessed ALS-data interactively
- In Austria, ALS-data will cover the whole country by the end of 2012
- DSM from photogrammetry will be available every 5 years.
- Multi-temporal DSM's will allow accurate monitoring methods for forest growth and felling activities.

2 Operating figures describing manageable forest units

Umweltdata is not a research institute. We offer services in the field of FMP, GIS and Remote Sensing (RS). We are trying to apply new methods in order to
- obtain better results (more significant, more accurate)
- provide faster delivery to our customers
- minimize the costs.

In any case we are driven by the needs of our customers. They want to get operating figures describing manageable forest units in order to run their daily business. And their forest areas look like FuzzyForests in many cases.

Table 1: Commonly used properties describing forest stands (sub-departments)

Parameter	Acronym	Unit	Reliability of RS-data			Source	Priority in Planning and Management (1=high, 4=low)
			coniferous	decidous	mixed		
Tree Height	H	m	++	++	++	ALS-data	2
Diameter	DBH	cm	+/-	-	-	ALS-data	2
Basal Area	BA	m2/ha	+/-	-	-	ALS-data	4
Stock Volume	SV	m3/ha	+	+/-	+/-	ALS-data	1
Tree Species	TS		+/-	+/-	+/-	CIR-Image	1
Age	t	y	0	0	0	local records	3
Number of Trees	NT	1/ha	+	+/-	+/-	ALS-data	2
Canopy Coverage	CC	percent	++	++	++	ALS-data	1
Stock Density	SD	percent	+	+/-	+/-	ALS-data	3
Growth Rate	GR	m3/ha/y	+	+	+	multi-temporal ALS	3
Felling Maturity	could be deduced from Tree Height, Canopy Coverage and Stock Volume						2
Thinning Priority	could be deduced from Tree Height, Canopy Coverage and Number of Trees						2
Harvesting Technology	could be deduced from the DTM						2

In table 1, the listed reliability is a rough estimation based on nationwide available ALS-data with 2-4 pulses per m^2 and the methods, mainly published by Markus Hollaus (Hollaus et al. 2009, 2010, 2011). In the case of tree species, I refer to promising results with WorldView2-data (Immitzer et al, 2011).

If data originating from a grid-based inventory is available, we use it to calibrate the results from RS. During the inventory fieldwork, we put a high effort in to GPS-measurements on the sample plots in order to co-reference inventory- with RS-data properly.

The most important figures (H, SV, TS, CC, SD) are combined into a continuous raster model. We propose a ground resolution of 5 meters; 2 meters or 10 meters would be suitable as well. Most figures are smoothed by focal functions with focal radiuses of 5 to 20 meters. Sometimes the results refer as mean-values to existing vector polygons like sub-departments or land parcels.

3 Maps supporting forest management

Until the proliferation of GIS in forestry, forest managers had to be satisfied with one single type of thematic map. In Austria, this map (scale 1:10.000) showed age classes in most cases, sometimes ortho-images. The diversity of themes coming from ALS chased us into multiple map books with two, sometimes four or even six different subjects combined into one volume. Some examples are presented in figures 6 to 10.

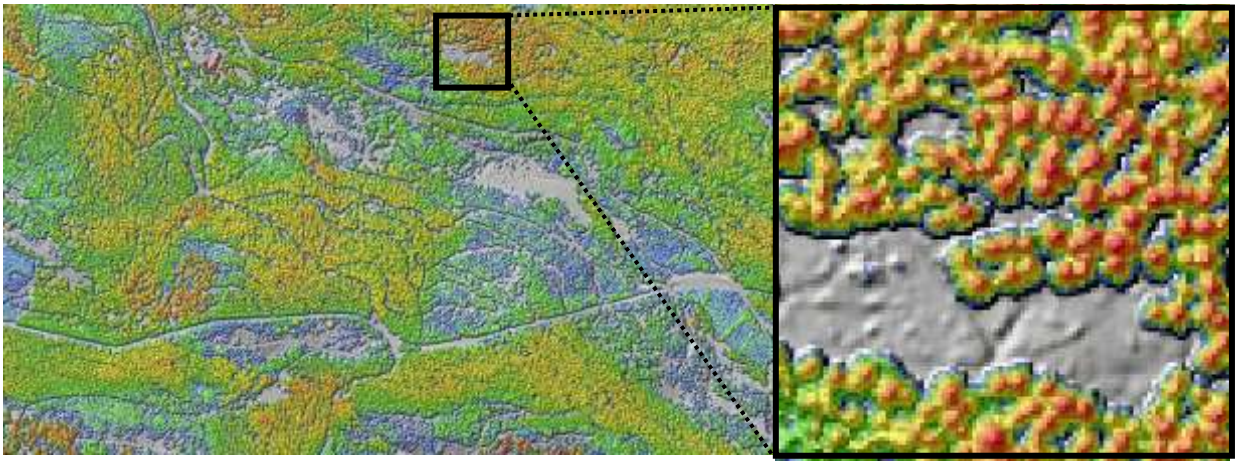


Figure 6: fusion of canopy shaded relief with colors from canopy elevation (legend see figure 3)

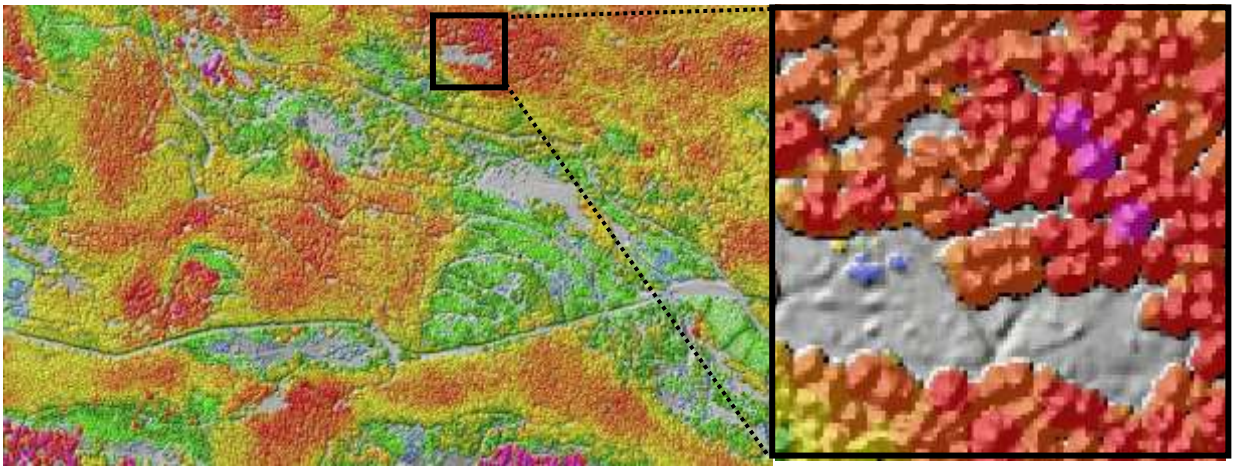


Figure 7: fusion of canopy shaded relief with colors from tree heights (legend see figure 3)

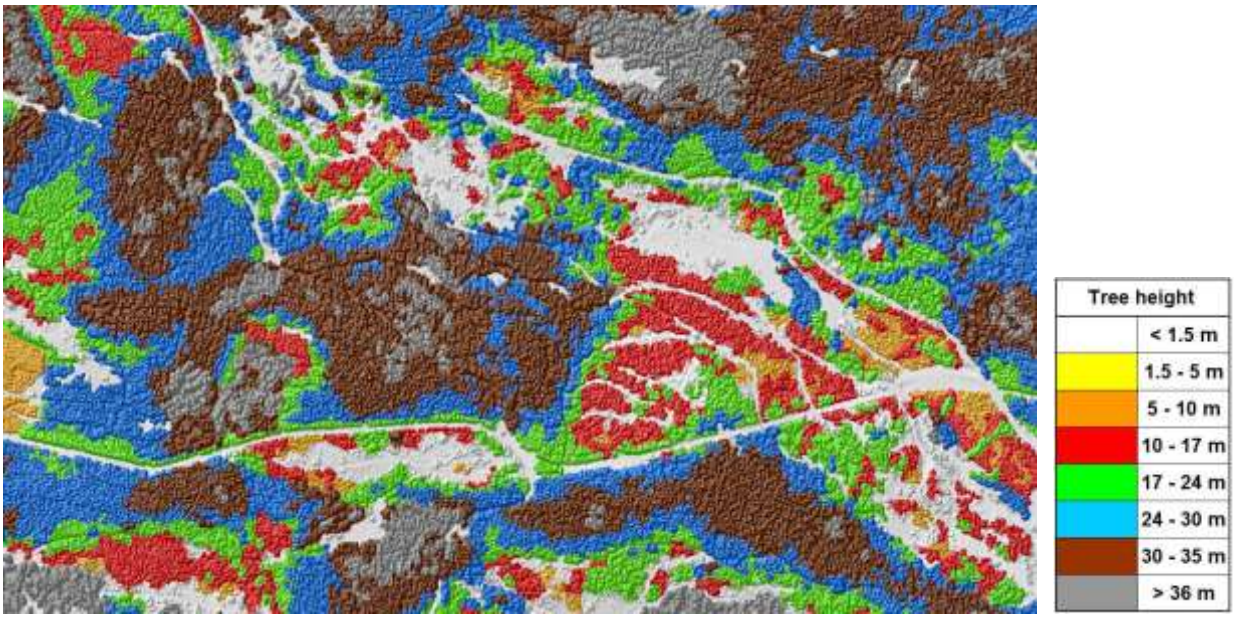


Figure 8: fusion of canopy shaded relief with generalized tree height classes

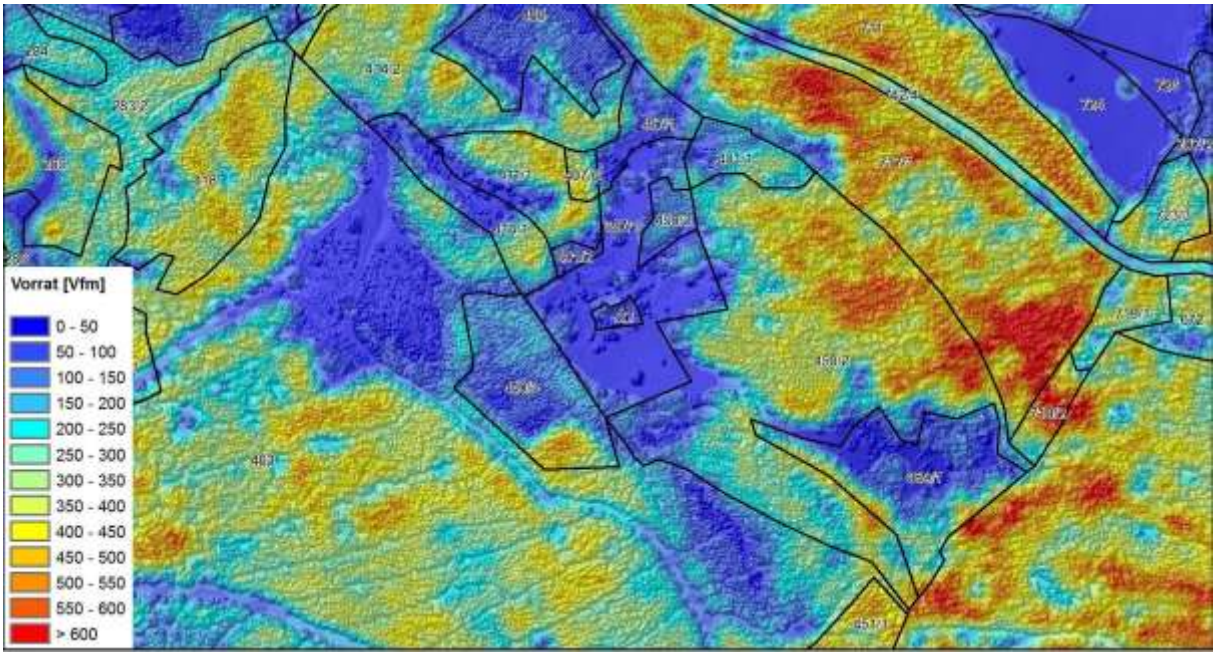


Figure 9: Stock volume estimation combined with land parcels

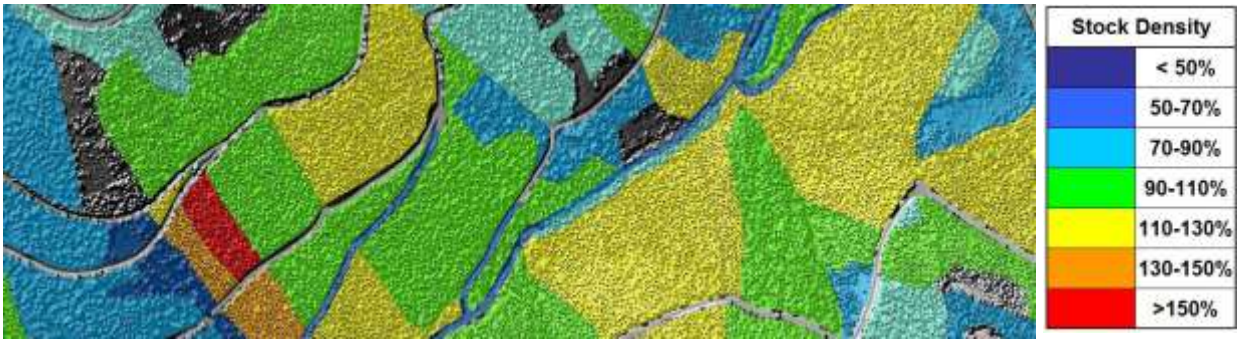


Figure 10: Stock density dependent from tree height as indicator for thinning priority

4 Prospects: future planning processes

Spec's for the ultimate planning tool in FuzzyForests:

- Running autonomously on a tablet PC without internet access
- Replicating automatically to a central server if internet access is available
- Accessing interactively to a continuous raster models of the forest
- Calculating expected timber volume
- Giving feedback by sustainability indices

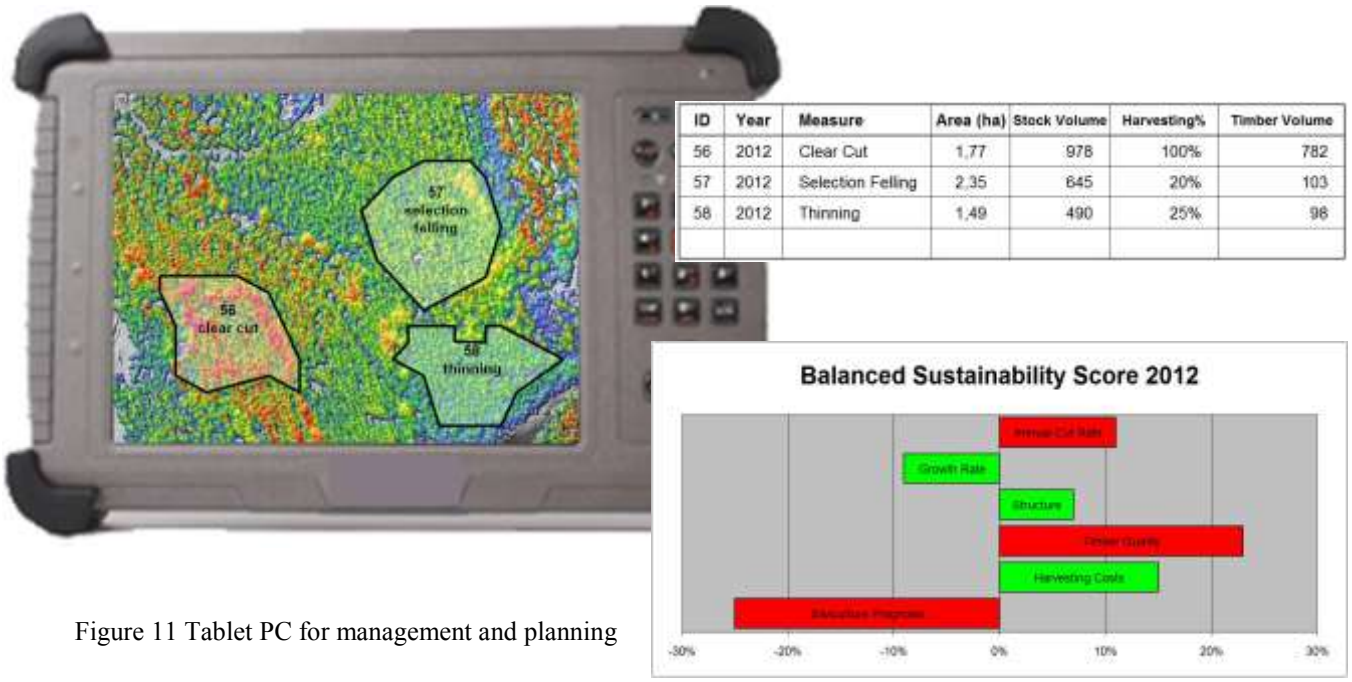


Figure 11 Tablet PC for management and planning

References

Hollaus, M., Wagner, W., Schadauer, K., Maier, B., Gabler, K., 2009. Growing stock estimation for alpine forests in Austria: a robust lidar-based approach. *Canadian Journal of Forest Research* 39(7), 1387-1400.

Hollaus M., Eysn L., Mücke W., Pfeifer N., Mandlburger G., 2011. Forest delineation and structure assessment based on airborne laserscanning data. Proceedings of European LiDAR Mapping Forum, Salzburg, November 2011

Immitzer M., Koukal T., 2011 Möglichkeiten der Baumartenerkennung auf Basis von Worldview-2 Satellitenbildern. Feasibility Study for Umweltdata GmbH

specialists would deny the need of segmented small units. With their statistical approach, they can perfectly answer the questions of stock volume, annual growth, annual cut rates and long-term sustainability figures for larger units.

On the other hand, the forest ranger wants to get rid of some manageable units. And he will need maps for his day-to-day business.

In many European regions, the commercial forestry has developed from a clear-cut-based tradition to different kinds of selection-felling-methods. The application of such nature-near techniques results in â€˜Fuzzy Forestsâ€™ after some decades, forest areas, which will persistently resist any kind of vector segmentation.

Fuzzy Forests are a challenge in FMPâ€™s as well as in forest cartography. After realizing, that splitting Fuzzy Forests into small segments means barking up the wrong tree, we came to the conclusion, that Lidar-based forest assessment will perfectly meet our requirements.

After mapping 300.000 hectares of forest land with airborne LIDAR data, I will present different kinds of successfully proven thematic maps (fuzzy or segmented), discuss the methods used in production, highlight the usability and the feedback from forest rangers and give an outlook, how such geo-data will generate added-value as interactive FMPâ€™s on tablet PCâ€™s.

Operational forest inventory and planning using airborne LiDAR in Tasmania, Australia

Dr Robert Musk[1], Timothy Osborn[2] & David Mannes[3]

[1]Forestry Tasmania rob.musk@forestrytas.com.au
[2] Forestry Tasmania tim.osborn@forestrytas.com.au
[3] Forestry Tasmania david.mannes@forestrytas.com.au

Paper Number: SL2012-200

1. Introduction

Small footprint airborne LiDAR systems generate high resolution point clouds that represent three dimensional forest canopy structure and so may be used to inform forest inventories (e.g. Evans et al., 2006; Lefsky et al., 1999; Lim et al., 2003; Magnussen & Boudewyn, 1998; Maltamo et al., 2004; Næsset, 1997a; Næsset, 1997b; Næsset, 2002; Nelson et al., 1988; Reutebuch et al., 2005; Ritchie et al., 1993). Forestry Tasmania, as the agency responsible for the management of Tasmanian state forests, has commenced acquisition of small footprint discrete return LiDAR data over its entire estate. Amongst several other objectives, the organisation seeks to improve the precision of its wood inventory using this dataset. Embedded within the estate is a plantation resource of *Eucalyptus nitens* and *E. globulus* covering 42 000 Hectares that is predominantly managed for a high value sawlog yield. Most of this plantation comprises first rotation plantings on sites formerly carrying native forest. The spatial variability in stand condition and woody weed load is very high, as is the demand for precise wood inventory. This work follows that of numerous others in deriving raster-based LiDAR variables, such as height quantiles and proportion of returns from within, or below, specific height classes, which are then used to develop a model to impute and map stand characteristics in the plantation resource. Careful application of feature selection procedures is employed to ensure model robustness. The large amount of data that the inventory produces has led us to develop a new software tool for use by forest planners that allows them to access and query the data. Its use has led to improved operational planning outcomes.

2. Methods

2.1 Study area

The study area comprises Forestry Tasmania's eucalypt hardwood plantation estate covering approximately 24 000 Ha in the east of Tasmania centred about 42°S 148° E. Figure 1 presents a map of Tasmania showing the extent of the State Forest managed by Forestry Tasmania and the plantation resource embedded within it. To date, the LiDAR coverage extends throughout the east and south of the estate pictured. These plantations are managed for a variety of harvest products, but the principle aim is to produce high value, pruned solid wood. To this end, most stands are managed over a 22-25 year rotation with three pruning lifts on selected trees and a single thinning operation between 5 and 12 years of age. Management decisions regarding the scheduling, and within-compartment extent of pruning, thinning and harvesting operations rely on access to precise stand information.

2.2 Field data

Field plots were established using a stratified random sampling schema across the component of the plantation eucalypt estate within the LiDAR data capture boundary for the summer campaigns from 2009/2010 to 2011/2012 (N = 253). Compartment age formed the basis of the strata. The mean time difference between field plot measurement and LIDAR data capture was 2.1 months. The maximum time difference was 7.3 months. Field plots were circular with a 12.0m radius (452 m^2). Field plot

centres were geolocated using a Trimble Pro-XR GPS unit. At least 200 GPS points were recorded at each plot and these were differentially corrected in a post processing step that ensured sub-metre planimetric standard errors. On each field plot, the magnetic bearing and horizontal distance to all trees with stem overbark diameter at 1.3m above ground (DBH) over 5cm was recorded. The tree species, total height (TH) and DBH of all trees over 5cm were recorded. All tree DBH were measured using diameter tapes. A Vertex ultrasonic hypsometer was used to measure total tree heights and tree distances from the plot centre. Ocular estimates of average canopy height and percentage canopy cover for all understory species with estimated canopy cover over 10% were also recorded. Three photos were taken from the plot centre facing north, south and vertically. These provided a qualitative record of field plot conditions.

Mean dominant height (MDH) was estimated as the weighted average of field measured tree heights (50 tallest trees Ha^{-1}). Bark thickness was estimated from unpublished Forestry Tasmania models. Basal area was calculated using underbark DBH and stocking (SPH).

2.2 LiDAR data

AAM acquired the LiDAR data during the summer months of 2009-2010 to 2011-2012 using an Optech Gemini discrete-return scanner. The maximum scan angle was set at 20° from nadir and the minimum point density was 200 points per 10 square meters. Up to four returns were recorded per pulse. The data were classified as ground or non-ground using proprietary algorithms and delivered in LAS 1.2 format. Each LAS file comprised a 600×600 metre tile with a 50 metre overlap to tiles adjacent. Vegetation height was calculated after interpolating a 1 metre resolution digital terrain model from the ground-classified returns using a multi-level B-spline (Lee et al. 1997).

LiDAR returns with a 5 metre buffer that were spatially coincident to the field plots were extracted from the delivered data and range of candidate predictor variables then calculated by means of two separate processes. In the first process, the returns were clipped to the extents of the field plot boundaries and numerous grid-based candidate predictor variables were then extracted. These comprised percentiles and moments of both vegetation height and intensity, proportions of returns within vegetation height and intensity strata defined in both absolute and relative terms. In the second process, a 0.5 metre resolution canopy height model (CHM) was interpolated from the vegetation height data for each field plot. The CHMs were then clipped to the extents of the field plot boundaries and equivalent pixel, rather than return metrics were then calculated.

A complete list of the candidate predictor variables derived through these two processing steps appears in Table 1 of the Appendix. Several potential predictors were not used in this study. In particular, the timing and intensity of any pruning and/or thinning operations was ignored since the spatial accuracy of operational boundaries is not sufficient for high resolution mapping purposes. In total, 236 candidate predictor variables were available for modelling.

2.3 Modelling strategy

Our task was to impute the tree lists obtained in the field observations from the LiDAR data. One consequence of deriving such a large list of candidate explanatory variables from the LiDAR data is that it is likely to contain numerous spurious variables. In order to identify the best predictor variables the imputation was achieved in a two stage process.

In the first stage, a feature selection procedure was applied to identify the candidate explanatory variables from the LiDAR data set that were most likely to have responded to the field observed data. This feature selection procedure entailed modelling a range of separate continuous response variables derived from the tree lists using two data mining techniques: tree bagging (Peters, et al., 2002) and random forests (Breiman 2001). The performance of the models was inspected. Where models demonstrated efficacy in describing the continuous response variables, the list of explanatory variables contributing to the model was recorded.

Having generated a reduced list of candidate variables through the feature selection procedure, we then undertook a forward stepwise model building approach to identify a set of candidate imputation models. The data was randomly split into fit (N = 190) and prediction sets (N = 63). The random forests algorithm was used to impute the field plots. This process was repeated with random data splits 100 times to generate 100 candidate imputation models. These candidate models were assessed in terms of their performance in predicting the continuous response variables described above using a subsequent 5-fold cross validation model refit. We used four criteria for model assessment.

Mean deviation of the residuals, a measure of bias, was defined by:

$$MD = \frac{\sum_{i=1}^{n} y - \hat{y}}{n} \quad (1)$$

and:

$$\%MD = \frac{\sum_{i=1}^{n} 100\left(\frac{y - \hat{y}}{y}\right)}{n} \quad (2)$$

Mean absolute deviation of the residuals, a measure of precision, was defined by:

$$MAD = \frac{\sum_{i=1}^{n} abs\left(y - \hat{y}\right)}{n} \quad (3)$$

and:

$$\%MAD = \frac{\sum_{i=1}^{n} 100\left(abs\left(\frac{y - \hat{y}}{y}\right)\right)}{n} \quad (4)$$

All candidate predictor variables, together with an imputed plot identifier, selected in the optimal imputation model were then derived over the study area at a 4 metre cell resolution.

All LiDAR processing and modelling work was undertaken using open-source software. The principle software tool used was the R language and environment (R Development Core Team 2011) with the additional packages: sp (Pebesma and Bivand, 2005; Bivand, Pebesma and Gomez-Rubio, 2008), MBA (Finley and Banerjee, 2010), data.table (Dowle and Short, 2011), moments (Komsta and Novomestky, 2007), rgdal (Keitt et al., 2010), caret (Kuhn et al., 2012), ipred (Peters and Hothorn, 2012), yaImpute (Crookston and Finley, 2007), foreach (Revolution Analytics, 2011) and doRedis (Lewis, 2012). The high computing load involved in generating large area rasters was accommodated using 50 Xeon cores in a virtual machine cluster that was managed using redis (Lerner 2010) on an openSUSE 11.0 Linux operating system.

3. Results

3.1 Field data description

Scatterplots of the first two L-moments of tree height (TH-L1, TH-L2) and DBH (DBH-L1, DBH-L2) together with MDH, BA and SPH appear in Figure 2. Strong relationships are apparent between basal area and mean dominant height and between the first L-moments of tree height and DBH. Stocking varies widely and is quite high is some plots, reflecting the presence of double leader stems in many areas. Basal area and stocking show skew that is typical in these stands. The tree metrics are also skewed and consequently the first two L-moments are correlated.

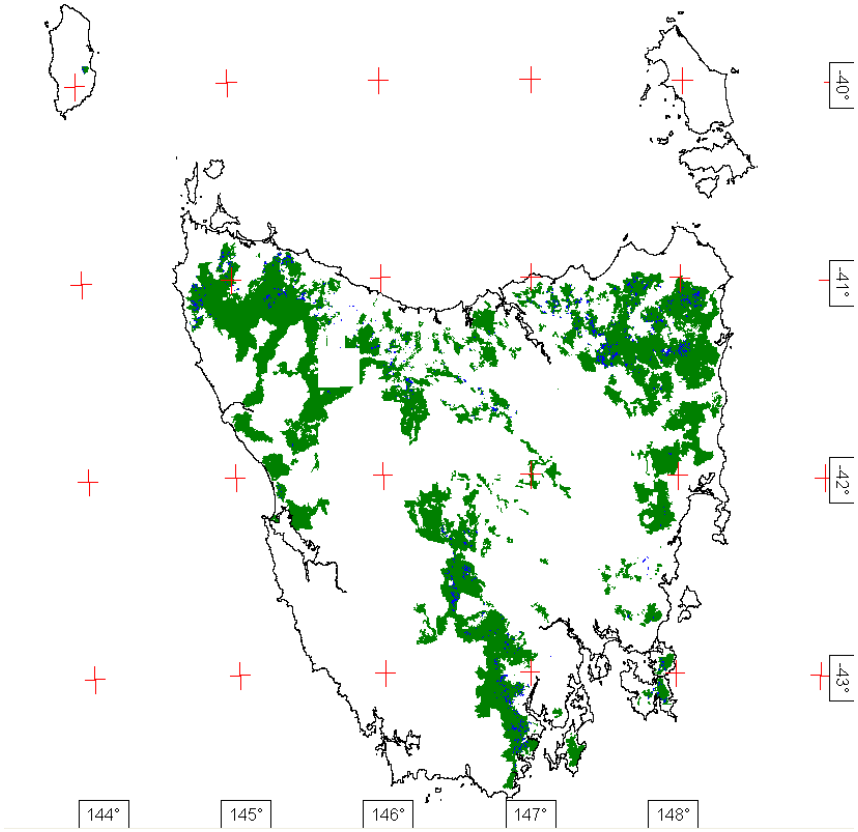


Figure 1 Map of Tasmania, Australia, showing State forest (green) with embedded eucalypt plantation (blue) managed by Forestry Tasmania. LiDar coverage to date extends across the eastern half of the estate.

3.2 Feature selection

Table 1 presents model R^2 values resulting from the two feature selection approaches: tree bagging and random forests. It is evident from the results in this table that the LiDAR variables are very weak predictors of the third and fourth L-moments of tree height and DBH. As a result, the predictor variables for these models were not included in the subsequent stage. The two approaches generate very similar model R^2 values and selected similar predictor variables (not shown), providing some comfort that the procedures were behaving as expected. Table 1 in the Appendix presents the reduced list of candidate predictor variables. The reduced list is dominated by pixel metrics, suggesting there is no information content within the canopy that cannot be attained from the canopy surface. Very few intensity metrics remain. This is somewhat surprising, as earlier work with a reduced subset of these data had suggested that the return intensity could be used to discriminate crop and non-crop species in these plantations.

Table 1 Dependent variable and model R^2 resulting from two feature selection approaches: tree bagging and random forests

Dependent variable	Tree bagging model R^2	Random forest model R^2
H1	0.893	0.912
H2	0.426	0.504
H3	0.132	0.157
H4	0.025	0.037
D1	0.731	0.758
D2	0.486	0.583
D3	0.087	0.077
D4	0.047	0.034
BA	0.567	0.592
MDH	0.931	0.950
SPH	0.300	0.398

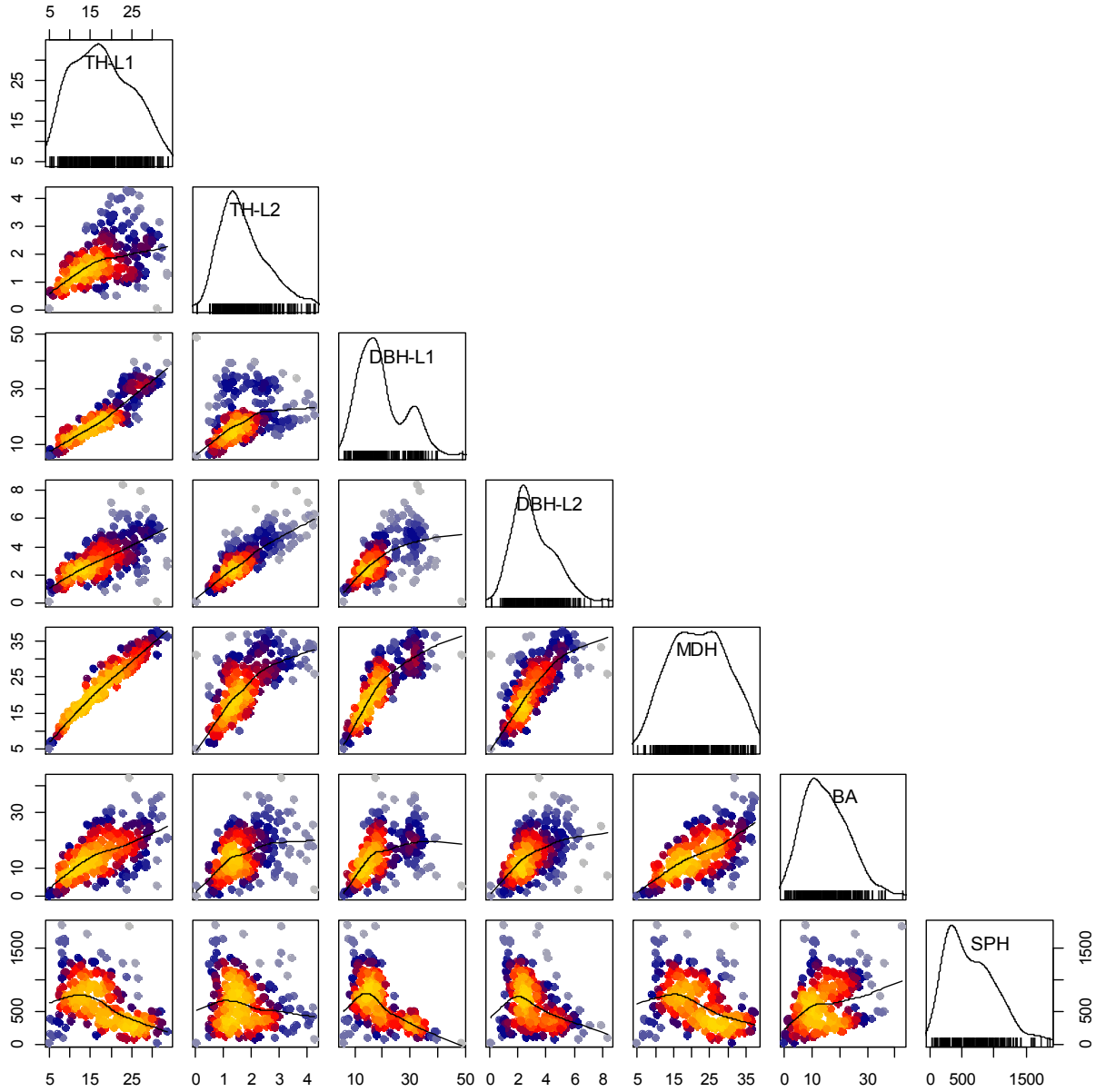


Figure 2 Scatterplots matrices (with spline curves) of field plot data: the first two L-moments of tree height and DBH, together with MDH, BA and SPH. Scatterplot colors represent density of observations from yellow (high density) to grey (low density).

3.3 Imputation

Scatterplots of the criteria statistics generated by the iterative fitting process were inspected. A subset example is presented in Figure 3 for illustrative purposes. There were consistently positive correlations between % MD and % MAD statistic values, and between MD and MAD statistic values (not shown) and consistently strong positive correlations between dependent variables. These results indicated a strong preferential ranking of candidate models.

In order to choose the optimal model composite weighted %MD and %MAD statistics were derived in the final step. The composite statistics were the mean statistic values across each variable, weighted by the inverse of the statistics absolute value. This weighting ensured that models that performed comparatively well for strongly predicted metrics (e.g. TH-L1) were favoured over models that preformed comparatively well for weakly predicted metrics (e.g. SPH). Figure 4 shows the scatterplot between composite %MD and %MAD statistics for each of the iteratively fitted imputation models.

The strong correlation ensures a clear choice can be made for a final model weighing up the preference for precision versus bias. We chose model 37 as the optimal model. Fit statistics for the optimal imputation model appear in Table 2. Figure 5 presents scatter plots of observed versus predicted field metrics. The model performs well in predicting the 1st L-moments and MDH. It performs less well in predicting 2nd L-moments and BA. Other metrics are poorly predicted.

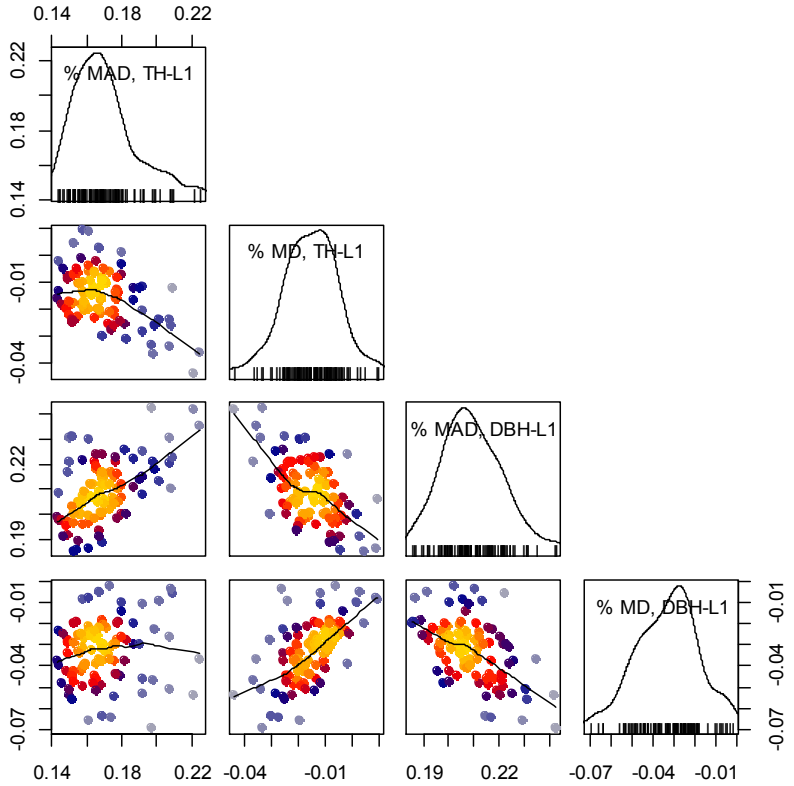


Figure 3 Density and scatterplots (with spline curves) of %MAD versus %MD from the 1st L-moments of tree height and DBH for the iteratively fitted imputation models. Scatterplot colors represent density of observations from yellow (high density) to grey (low density).

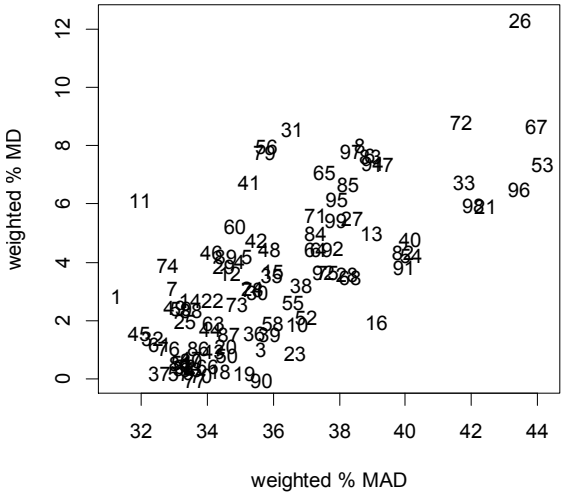


Figure 4 Composite weighted %MAD versus %MD for the iteratively fitted imputation models. Model 37 was identified as the optimal model.

Table 2 Fit statistics for imputation model 37.

Variable	H1	H2	H3	D1	D2	D3	MDH	BA	SPH
MD	0.34	-0.04	0.00	0.66	-0.02	-0.01	0.20	-0.27	-26.67
MAD	2.26	0.63	0.29	3.68	0.94	0.46	2.15	5.78	317.48
% MD	0.75	-33.00	212.85	-0.03	-37.01	51.59	0.31	-27.67	-52.80
% MAD	13.94	61.62	794.41	18.46	63.56	387.97	11.48	61.06	87.31

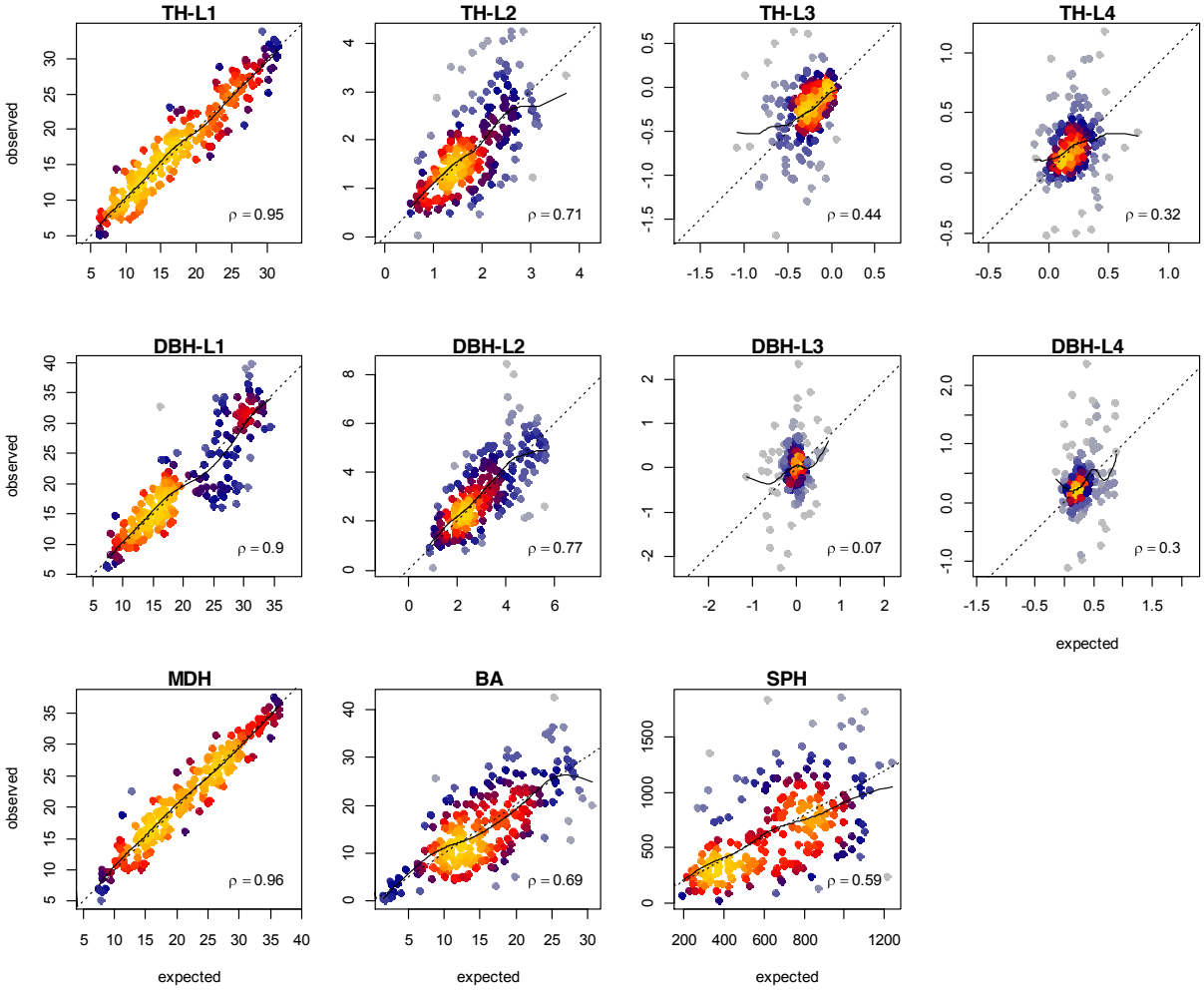


Figure 5 Observed versus expected stand metrics from the imputation model 37, together with spearmans rho correlations, 1:1 lines (dotted) and spline curves (solid). Colors represent density of observations from yellow (high density) to grey (low density).

3.4 Implementation

The capacity of LiDAR-based inventory to provide within-compartment information is a key advantage of the approach over traditional field-based inventories. In order to realise the full potential of the data we developed a simple database structure and a software program that allows operational forest planners to query and view the imputed inventory. The imputation data is stored in an Oracle 10.1.05 database accessible to all staff via an intranet. A points table contains each imputed pixel,

geo-referenced with an easting and northing. The values are encoded as an Oracle SDO_GEOMETRY datatype. Each unique imputation has a yield table entry associated with it in a yield table. Currently the yields for each product type associated with each imputed pixel are pre-generated for each asset using Forestry Tasmania' growth and yield system.

The software program is designed to retrieve and display the data stored in these two tables based on spatial overlays with an operations database that contains spatial, and associated silvicultural asset data. The software program is written in C# / .NET 3.5. The user selects an asset of interest in a GUI interface. In doing so, the points table is spatially overlaid onto the asset polygon shape in order to extract all imputed pixels that fall within the asset boundary. The yield information for those imputed pixels is then retrieved from the yield table and stored on the local machine in a generic list.

The user is presented with a main form which contains a TreeView showing a District / Block / Asset hierarchy. The top left of the form provides a list of all assets for which LiDAR inventory is available. Selecting an asset populates the centre top panel with asset data (such as asset name, silvicultural prescription and planting year) and the right top panel with a thumbnail view of the asset shape with proximate roads and streams. Once an asset has been chosen the user selects the 'Load Lidar' button and this executes the spatial overlay for the asset and loads all the imputed pixels as described above. In the lower half of the form summary charts and tables of product volumes yields by year are then presented, together with a detailed map of the Asset in a second, separate form. In this separate form the map shows each imputed pixel colour-coded according to the projected volume of the selected product in a selected year. A slider bar allows the user to display this product map for any year on the planning horizon. There's also a 'Rescale Colours' option that allows the user to amplify the colour range by observing the limits (i.e. min, max) from the current year, rather than the limits from the entire planning horizon.

A typical compartment (asset) contains around 5 million imputed pixels. It takes an average of 200 seconds to load the generic list. Each operation on the list (e.g. search for a particular year and sum of products, and paint a screen map) takes an average of 1.5 seconds. Currently we use a simple linear search to retrieve the data from the Oracle database and there is scope to improve performance with some form of indexing.

4. Discussion

Imputation based LiDAR inventories allow projection of the field observed tree lists using legacy growth and yield systems and so provide a temporal view of stand characteristics. The information provided has utility across all tiers of planning, and serves to unify operational, tactical and strategic planning, providing consistent estimates into each. The principal limiting factor in using LiDAR based inventories is no longer the age of the underlying LiDAR data, but the quality of the imputation and the capacity of an organisations growth and volume models to accurately project plot data forward in time.

The utility of the imputed data is demonstrated by the range of uses it is put to by Forestry Tasmania. Using the tool described above, information derived from the imputation is accessed and interrogated. Its principal use is for operational decision support. The tool allows visual representation of the harvesting unit in spatial and temporal dimensions. The high resolution of the imputed data facilitates planning of operations at a within-coupe level. Imputed estimates of harvest volume and stem size distribution provide insights into harvesting economics and can inform the selection of appropriate harvesting equipment. Improving inventory precision also reduces the required risk loading in harvest rates. As a tactical planning tool yield projection provides information that informs silvicultural planning and scheduling. In particular, disjunctive decisions such as when to thin a stand to minimise windthrow risk are supported by tree list based yield projection. The underlying yield tables are also useful for strategic, estate wide planning exercises with the added advantage of being directly informed by the same estimates as the operational and tactical planning processes.

References

Bivand, R.S. , Pebesma, E.J. And V. Gomez-Rubio (2008) *Applied spatial data analysis with R*. Springer, NY. http://www.asdar-book.org/

Breiman, L. 2001. Random forests. *Machine Learning* 45:15–32.

Coops, N.C., Hilker, T., Wulder, M.A., St-Onge, B., Newnham, G., Siggins, A., Trofymow, J.A., 2007. Estimating canopy structure of Douglas-fir forest stands from discrete return LiDAR. *Trees* 21, 295–310.

Crookston, N.L and Andrew O. Finley, A.O. (2007) yaImpute: An R Package for kNN Imputation. *Journal of Statistical Software*, 23(10): 1548-7660

Cutler, R. D., Edwards, T. C., Jr., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., et al. (2007). Random Forests for classification in ecology. *Ecology*, 88, 2783−2792.

Dowle, M. and T. Short. (2011). *data.table: Extension of data.frame for fast subset, fast merge and fast grouping*. R package version 1.5.3. http://CRAN.R-project.org/package=data.table

Evans, D. L., S. D. Roberts and R. C. Parker (2006). LiDAR – A new tool for forest measurements? *The Forestry Chronicle* 82(2): 211-218.

Falkowski, M.J., Evans, J.S., Martinuzzi, S., Gessler, P.E. and A.J. Hudak Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment* 113 (2009) 946–956

Finley, A.O. and S. Banerjee (2010). *MBA: Multilevel B-spline Approximation*. R package version 0.0-7. http://CRAN.R-project.org/package=MBA

Gabriel, E., Fagg, G.E., Bosilca, G., Angskun, T., Dongarra, J.J., Squyres, J.M., Sahay, V., Kambadur, P., Barrett, B., Lumsdaine, A., Castain, R.H., Daniel, D.J., Graham, R.L., and T.S. Woodall (2004) Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation. In *Proceedings, 11th European PVM/MPI Users' Group Meeting*, Budapest, Hungary, September 2004.

Hosking, J. R. M. (1990). *L*-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B*, **52**: 105-124.

Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., & Falkowski, M. J. (2008). Nearest neighbor imputation modeling of species-level, plot-scale structural attributes from lidar data. *Remote Sensing of Environment*, 112, 2232−2245.

Keitt, T.H., Bivand, R.S., Pebesma, E.J. and B. Rowlingson (2010). *rgdal: Bindings for the Geospatial Data Abstraction Library*. R package version 0.6-28. http://CRAN.R-project.org/package=rgda

Komsta, L. and F. Novomestky (2007). *moments: Moments, cumulants, skewness, kurtosis and related tests*. R package version 0.11. http://www.r-project.org, http://www.komsta.net/

Kuhn, M. with Wing, J., Weston, S., Williams, A, Keefer C. and A. Engelhardt (2012). *caret: Classification and Regression Training*. R package version 5.15-023. http://CRAN.R-project.org/package=caret

Latifi, H., Nothdurft, A. And B. Koch (2010) Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. *Forestry* 83(4): 395-407

Lawrence, R. L., Wood, S. D., & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment*, 100, 356−362.

Lee, S., Wolberg, G. and S. Y. Shin (1997) Scattered data interpolation with multilevel B-splines *IEEE Transactions on Visualization and Computer Graphics*, 3(3):229-244

Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D., (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment* 70, 339–361.

Lerner, R.M. (2010) At the forge: Redis. *Linux Journal* 197(5): 1075-3583

Lewis, B. (2012). *doRedis: Foreach parallel adapter for the rredis package*. R package version 1.0.5. http://CRAN.R-project.org/package=doRedis

Lim, K., P. Treitz, K. Baldwin, I. Morrison and J. Green (2003). Lidar remote sensing of biophysical

properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing* 29(5): 658-678.

Magnussen, S. and P. Boudewyn (1998). Derivations of stand heights from airborne laser scanner data with canopy based quantile estimators. *Canadian Journal of Forest Research* 28(7): 1016-1031.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J., Vehmas, M., 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sensing of Environment*. 90, 319–330.

Næsset, E. (1997a). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment* 61(2): 246-253.

Næsset, E. (1997b). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* 52(2): 49-56.

Næsset, E., (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* 90, 88–99.

Næsset, E. and T. Økland, (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment* 79, 105–115.

Nelson, R., R. Swift and W. Krabill (1988). Using airborne lasers to estimate forest canopy and stand characteristics. *Journal of Forestry* 86(10): 31-38.

Pebesma, E.J. and R.S. Bivand (2005) Classes and methods for spatial data in R. *R News* 5 (2), http://cran.r-project.org/doc/Rnews/.

Peters, A. and Hothorn, T. (2012). *ipred: Improved Predictors*. R package version 0.8-13. http://CRAN.R-project.org/package=ipred

R Development Core Team (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.

Reutebuch, S. E., H. Andersen and R. J. McGaughey (2005). Light Detection and Ranging (LIDAR): An Emerging Tool for Multiple Resource Inventory. *Journal of Forestry* 103(6): 286 - 292.

Revolution Analytics (2011). *foreach: Foreach looping construct for R*. *R package* version 1.3.2. http://CRAN.R-project.org/package=foreach

Ritchie, J. C., D. L. Evans, D. Jacobs, J. H. Everitt and M. A. Weltz (1993). Measuring canopy structure with an airborne laser altimeter. *Transactions of the ASAE (American Society of Agricultural Engineers)* 36(4): 1235-1238.

Appendix

Table 1: Candidate predictor variables. Number derived from the LiDAR data and reduced number applied in modelling resulting from feature selection approaches: Tree bagging and random forests

Variable(s)		Number	
Name	Description	Derived	Applied
isvar, isskew, iskurt	Pixel intensity moments (variance, skewness, kurtosis)	3	1
brqsi0901, brqsi0902, ... , brqsi0201	Pixel intensity density strata (rqi09 – rqi01, rqi09 – rqi02, ... , rqi02 – rqi01)	36	4
bnqsi0901, bnqsi0902, ... , bnqsi0201	Pixel intensity strata (nqi09 – nqi01, nqi09 – nqi02, ... , nqi02 – nqi01)	36	0
rqsi01, rqsi02, ... , rqsi09	Pixel intensity density (proportion of Pixel with intensity higher than: 10%, 20%, ...:, 90% of maximum intensity)	9	0
nqsi01, nqsi02, ... , nqsi10	Pixel intensity percentiles (10%, 20%, ... , 100%)	10	0
hsvar, hsskew, hskurt	Pixel height moments (variance, skewness, kurtosis)	3	2
brqsh0901, brqsh0902, ... , brqsh0201	Pixel height density strata (rqh09 – rqh01, rqh09 – rqh02, ... , rqh02 – rqh01)	36	8
bnqsh0901, bnqsh0902, ... , bnqsh0201	Pixel height strata (nqh09 – nqh01, nqh09 – nqh02, ... , nqh02 – nqh01)	36	11
rqsh01, rqsh02, ... , rqsh09	Pixel height density (proportion of Pixel with height greater than: 10%, 20%, ...:, 90% of maximum height)	9	1
nqsh01, nqsh02, ... , nqsh10	Pixel height percentiles (10%, 20%, ... , 100%)	10	7
ivar, iskew, ikurt	Return intensity moments (variance, skewness, kurtosis)	3	2
brqi0901, brqi0902, ... , brqi0201	Return intensity density strata (rqi09 – rqi01, rqi09 – rqi02, ... , rqi02 – rqi01)	36	3
bnqi0901, bnqi0902, ... , bnqi0201	Return intensity strata (nqi09 – nqi01, nqi09 – nqi02, ... , nqi02 – nqi01)	36	0
rqi01, rqi02, ... , rqi09	Return intensity density (proportion of returns with intensity higher than: 10%, 20%, ...:, 90% of maximum intensity)	9	0
nqi01, nqi02, ... , nqi10	Return intensity percentiles (10%, 20%, ... , 100%)	10	0
hvar, hskew, hkurt	Return height moments (variance, skewness, kurtosis)	3	2
brqh0901, brqh0902, ... , brqh0201	Return height density strata (rqh09 – rqh01, rqh09 – rqh02, ... , rqh02 – rqh01)	36	3
bnqh0901, bnqh0902, ... , bnqh0201	Return height strata (nqh09 – nqh01, nqh09 – nqh02, ... , nqh02 – nqh01)	36	6
rqh01, rqh02, ... , rqh09	Return height density (proportion of returns with height greater than: 10%, 20%, ...:, 90% of maximum height)	9	0
nqh01, nqh02, ... , nqh10	Return height percentiles (10%, 20%, ... , 100%)	10	6

Paper Abstracts

ERROR BUDGETS FOR LIDAR DRIVEN INDIVIDUAL TREE GROWTH PROJECTIONS

George Z. Gertner[1]

[1]Professor Emeritus
Department of Natural Resources and Environmental Sciences
University of Illinois, Urbana-Champaign
Urbana, Illinois 61801
USA
gertner@illinois.edu

Paper Number: SL2012-007

Abstract

The US Forest Service Forest Vegetation Simulator (FVS) is being used to evaluate a variety global change scenarios. This is a distance independent single-tree growth model developed for projecting tree and forest development through time. In this paper, an error budget is developed for FVS projections, where initial models inputs are based on single-tree stem maps developed with small-footprint airborne lidar. An error budget shows the overall precision of estimates/predictions made with a system, divided according to different types of error sources within and outside of the system. In a comprehensive fashion, sources of uncertainties due to measurements, classification, sampling, parameter estimates, are accounted for in the lidar derived stem maps and within the FVS system. Spatially identifying the sources of uncertainties, modeling their accumulation and propagation, and finally, quantifying them locally on a tree basis and globally on a forest level are presented. Uncertainties in future forest responses due to uncertainties in projected global climatic change predictions that will also drive this type of forest model will be also discussed.

There are benefits to having an error budget for users of these models. The primary benefit is the acknowledgment that errors and uncertainties exist; and it is reported. When using the model, the user will know that there is uncertainty and will use the model with full knowledge of the uncertainty and the risk of making the correct predictions (or incorrect predictions). The second most important benefit is that in a systematic fashion the decision-maker will understand the sources of errors and uncertainties in the predictions.

Comparing forest stand characteristics predicted from digital aerial photogrammetry and airborne laser scanning

Gobakken, Terje, Næsset, Erik & Bollandsås, Ole Martin

Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, P.O. Box 5003, NO-1432 Ås, Norway. terje.gobakken@umb.no,

Paper Number: 012

Abstract

Airborne laser scanning (ALS) is considered the most accurate remote sensing technology for retrieval of detailed 3D information about tree canopies and other biophysical characteristics essential for forest resource inventories, such as mean tree height, basal area, and stem volume. Even though the cost of ALS data acquisition has decreased in the Nordic countries during recent years, it is still a substantial part of the total forest inventory cost. Historically, aerial images have been widely used in forest inventory because of their relatively low acquisition cost and good availability. As well as with ALS, it is also possible to derive 3D forest canopy information by means of aerial photogrammetry. With new improved software and hardware for digital photogrammetry this might be a cost effective alternative. Thus, the objective of the current study was to compare forest stand characteristics predicted using the area-based approach and data from digital photogrammetry and ALS. In total 154 georeferenced field sample plots were distributed systematically throughout a 1000 ha study area. The size of each plot was 200 m2. The study area was divided into three strata according to age class and site quality. Regression models for were fitted for mean tree height, basal area, and timber volume using various canopy height and canopy density metrics derived by means of matching of aerial images and small-footprint laser scanning as explanatory variables in the analysis. The ground sample distance was 17 cm for the images and the ALS pulse density was 5 points m-2. The digital terrain model was derived from the ALS data. For ALS the stratum-specific regression models explained 74–94%, 83–97%, and 86–98% and of the variability in ground-truth mean height, basal area, and volume, respectively. The corresponding regressions developed from data derived from images explained 77-91, 81-92, and 86-95%. When using the area-based approach and evaluating on 63 independent test-stands, the standard deviations of the differences between predicted and ground-truth mean tree height, basal area, and timber volume as percentage og ground-truth values were 5.5-6.7%, 9.6-12.7%, and 11.6-12.4% for ALS and 6.1-8.0%, 10.5-16.7%, and 12.4-18.3% for photogrammetry, respectively. The results obtained by ALS were slightly better than the results obtained from photogrammetry, especially in young forest. Even though the standard deviation of the differences between predicted and ground truth values were larger for photogrammetry compared to ALS, aerial images might still be an interesting data source because of the low acquisition cost. In addition, the spectral information from images is useful for predicting tree species which is challenging to predict when only using ALS.

A model-based approach for estimating tree height distribution using airborne laser scanning data with and without field measurement data

Lauri Mehtätalo[1], Anni Virolainen[2], Jukka Tuomela[3] & Petteri Packalén[4]

[1]University of Eastern Finland, School of Forest Sciences lauri.mehtatalo@uef.fi
[2]University of Eastern Finland, School of Forest Sciences anni.virolainen@uef.fi
[3] University of Eastern Finland, Department of Physics and Mathematics
jukka.tuomela@uef.fi
[4]Oregon State University, Department of Forest Engineering petteri.packalen@uef.fi

Paper Number: #SL2012-016

Abstract

The airborne laser scanning (ALS) data of a forest contains direct information of the canopy height. This study introduced a theoretical model, which expresses the observed heights of the canopy envelope as a function of stand density, the parameters of the tree height distribution, and the shape of the individual tree crown. The model included the assumption of rectangular grid pattern of tree locations. The model was applied with empirical data from a eucalyptus plantation in two situations: with ground-measured training data and without ground measurements. In the first situation the idea is to estimate the parameters specifying individual tree crowns using laser data from plots with known tree heights. This information is then used with another set of plots to estimate the distribution of tree heights. The height distribution was assumed to follow Weibull distribution. In second situation, both the parameters specifying individual tree crowns and the distribution of tree heights are estimated simultaneously based on ALS data only. The main interest was to estimate the parameters of plot-specific tree height distribution; the estimation was performed by using the maximum likelihood method. The model for individual tree crown shape was also introduced, where the penetration of the laser pulse into the canopy was taken into account. The accuracies of the mean height and dominant height were evaluated and compared to the reference method, i.e., regression method. The mean height with ALS and ground-measured data resulted in root mean square error (RMSE) of 1.4 m and dominant height RMSE of 0.8 m. The mean height with ALS data only resulted in RMSE of 2.9 m and dominant height RMSE of 0.9 m. In both cases, the estimates of dominant height were more accurate than with the reference method, but the estimates of mean height were less accurate. The inconsistency of the dominant height and mean height estimation suggests that the Weibull distribution may not be the best option for the tree height distribution. Future studies focus on improving the model, e.g. allowing randomness to the crown shape or penetration, which were assumed to be fixed in this study.

Field trials of a full-waveform terrestrial laser scanner to measure forest canopy dynamics

F.M. Danson[1], R.P. Armitage[1], R. Gaulton[2], O.T. Gunawan[1] & F.A.Ramirez[1]

[1]School of Environment and Life Sciences, University of Salford, Salford M5 4WT, UK
f.m.danson@salford.ac.uk
[2]School of Civil Engineering & Geosciences, Newcastle University,
Newcastle upon Tyne NE1 7RU, UK rachel.gaulton@ncl.ac.uk

Paper Number: SL2012-019

Abstract

This paper presents the results of a field trial of the Salford Advanced Laser Canopy Analyser (SALCA), a unique two-wavelength full-waveform terrestrial laser scanner, designed to characterize spatial and temporal variations in forest canopy biophysical properties. The instrument is deployed on a tripod below the forest canopy, and records up to 9 million waveforms with a range resolution of 150mm across a full hemisphere. Two laser wavelengths 1064 and 1545 nm, with co-aligned beams, are emitted, and the maximum measurement range is approximately 100 m. The instrument was developed to exploit the information on forest canopy structure and composition contained in full-waveform returns, and the application of dual wavelength measurements to allow separation of foliage and woody material. SALCA was deployed on a weekly basis from April to September 2011, at a range of needle-leaved evergreen and broadleaf deciduous forest plots at Delamere Forest, Cheshire, UK. The laser scanner measured changes in leaf and needle cover, and the dynamics of leaf expansion associated with spring green-up in the deciduous plots. Coincident hemispherical photographs were acquired for comparison with the SALCA data, and record the key phenological events including bud burst and leaf unfolding.

Data pre-processing has focussed on extracting multiple returns, and associated intensity, by applying an adaptive noise threshold and identifying all maxima in the waveform. An empirical range-intensity correction method was developed to model and correct range-dependent decreases in backscattered energy. This enabled a simple ratio (1064/1545 nm) to be calculated for each return and, after thresholding, these data were used to classify points as woody material or foliage. Application of a point-based method to measure directional gap fraction provided estimates of leaf area index (LAI) at each stand through the growing season, and direct comparison with LAI determined from hemispherical photography. LAI estimates from the hemispherical photography were strongly correlated with LAI estimates from the laser scanner in both evergreen and broadleaved deciduous stands, but they were consistently higher when both woody and foliage points were included. Gap fractions based only on returns classified as foliage were similar in magnitude for the needle-leaved stands but, for the broadleaved deciduous stands, where measurements were made from leaf-off to full leaf-on stage, LAI was significantly lower, since woody plant area index was accounted for. Further validation is required through destructive sampling of test canopies, and methods to measure and account for foliage clumping must be developed.

Requirements of a habitat specialist in Swiss mountain forests – an assessment of forest structure and composition using laser remote sensing and field data

Florian Zellweger[1], Felix Morsdorf[2], Veronika Braunisch[3,4] & Kurt Bollmann[1]

[1]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, 8903 Birmensdorf, Switzerland, florian.zellweger@wsl.ch, kurt.bollmann@wsl.ch
[2]Departement of Geography, University of Zurich-Irchel, 8057 Zurich, Switzerland felix.morsdorf@geo.uzh.ch
[3]Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland, veronika.braunisch@iee.unibe.ch
[4]Forest Research Institute of Baden-Württemberg, 79100 Freiburg, Germany

Paper ID: SL2012-025

Abstract

Species richness in forest ecosystems largely depends on habitat structure and composition. These attributes can be assessed in field surveys, however, such data often lacks in spatial extent. Remote sensing technologies such as light detection and ranging (LiDAR) provide alternative tools to quantify structural elements across relatively broad areas at a fine resolution. To study the habitat requirements of hazel grouse (*Bonasa bonasia*), an indicator species of structurally rich forest stands, we assessed the structure and composition of Swiss mountain forests over three biogeographical regions.

We designed a sample based field survey of forest structure and composition and a LiDAR based assessment of vertical and horizontal forest structures using a nationwide LiDAR dataset with a mean point density of 1.4 m^2. The dependent variable consisted of hazel grouse presence/absence data at a resolution of 1 km^2. Species distribution models were computed for both variable sets separately and in combination, using boosted regression trees, a statistical machine learning technique.

Model performance assessment based on explained deviance and AUC showed that the combined model performed best, with over 55% explained deviance in the observed data, followed by the field and LiDAR models. The field model revealed that hazel grouse favored evenly distributed, rich ground vegetation, optimally with a substantial portion of bilberry (*Vaccinium myrtillus*). The abundance of tall rowans (*Sorbus aucuparia*), basal branched trees and a high percentage of resource trees in the shrub layer were found to be further essential habitat elements. LiDAR was powerful in detecting important structural features, whereby the horizontal forest structure explained more of the deviance than the vertical forest structure. The most influential LiDAR variable was a measure of canopy height heterogeneity. Apart from indicating structurally rich forest stands, it probably also served as a proxy of compositional aspects such as the abundance of light demanding resource trees and shrubs or of a well developed ground vegetation. To support habitat management, we derived variable thresholds at a relevant spatial scale (1 km^2) for forest management.

Our study showed that LiDAR provides adequate means to assess structural habitat elements area-wide, thus overcoming the difficulties associated with sample based field assessments. The best model fit, however, was obtained by combining LiDAR variables with compositional variables from the field survey. Hence, we successfully bridged the gap between different ecologically relevant scales, such as habitat configuration and structure at the regional scale and the abundance of habitat elements at the local scale. The methods applied in this study can also be used to identify hotspots of forest structural richness, a matter of interest in the light of emerging attempts to conserve biodiversity in forests.

The ICESat-2 mission and its potential for vegetation science

Amy Neuenschwander[1], Ross Nelson[2], Michael Lefsky[3]

[1]Applied Research laboratories, University of Texas at Austin
[2]NASA Goddard Space Flight Center
[2]College of Natural Resources, Colorado State University

Paper Number SL2012-030

Abstract

Plans are underway at NASA to launch the ICESat-2 (Ice, Cloud, and land Elevation Satellite) laser altimeter in 2016. While ICESat-2 is primarily designed to monitor changes in the cryosphere, it will also collect data over much of the Earth's vegetated surfaces. ICESat-2 will be a 532 nm photon-counting, multi-beam profiling, laser ranging system that requires not only new technology to collect the measurements but also new analysis techniques to extract information from the data. The combination of the laser repetition rate and satellite velocity will result in one outgoing laser pulse approximately every 70 cm on the Earth's surface. This laser repetition rate will allow detected photons to be accumulated in the along-track direction to recover surface elevation. Each outgoing laser beam is split into three pairs approximately 3 km apart on the Earth's surface (6 km total); each beam footprint on the surface is 10 m in diameter. To compensate for varying surface reflectance, each bear pair will consist of a strong beam (100 μJ) and a weak beam (25 μJ). The beam configuration as proposed for ICESat-2 is beneficial for terrestrial ecosystems as it enables a dense spatial sampling over the mid-latitudes (approximately 60°S – 65°N). To achieve a spatial sampling goal of no more than 2 km between equatorial ground tracks, ICESat-2 will be off-nadir pointed a maximum of 1.8 degrees from the reference ground track.

Within any one of the individual 10 m laser footprint, photons are detected directly before, during, and after the expected pulse returns (equivalent to above, within, and below the expected elevation of the pulse return) to capture all possible surface returns. For vegetated surfaces, the number of detected photons per strong beam laser shot is estimated to range between 0 to 4 photons, however more photons could be returned from tall vegetation. Solar background noise will be a significant challenge in the analysis of photon counting laser data as there is no way to distinguish green photons emitted by ICESat-2 (i.e. signal photons) from green photons returned from the atmosphere or scattered from adjacent targets (i.e. noise photons). New algorithms are being developed by the ICESat-2 Science Definition Team which will attempt to recover the ground and canopy surfaces for a direct retrieval of canopy height. Although the number of signal photons over vegetation is not expected to be high, there is potential for new, global digital elevation and canopy height models.

An industrial cost-benefit comparison of a LiDAR derived forest inventory to a traditional inventory for an annual harvest period

Murray Woods[1], Sebastien Lacroix[2], Chad St.Amand[3], Lino Morandin[4] and Doug G. Pitt[5]

[1]Ontario Ministry of Natural Resources, North Bay, Canada murray.woods@ontario.ca
[2]FPInnovations, Pointe-Claire, Canada, sebastien.lacroix@fpinnovations.ca
[3]Tembec, Timmins, Canada, chad.st-amand@tembec.com
[4]Tembec, Timmins, Canada, lino.morandin@tembec.com
[5]Canadian Wood Fibre Center, CFS, Sault Ste. Marie, Canada, dpitt@NRCan.gc.ca

Paper Number : SL2012-033

Abstract

Airborne LiDAR-derived 400-m^2 rasters of forest inventory attributes (heights, basal area, gross total volume, gross merchantable volume, DBHq, and biomass) were produced for a 630,000 ha active forest management unit near Timmins, Ontario, Canada, as supplements to the existing basic forest resource inventory. The LiDAR data set was wall-to-wall, at an average of 0.5 pulses per m^2 (Woods et al., 2011). A retrospective cost-benefit analysis was conducted on an annual harvest schedule to quantify the savings associated with the use of these LiDAR-derived inventory supplements. The real costs of more than 30 activities related to inventory acquisition and processing, forest operations, and mill processing, born from traditional knowledge, were compared to known industrial experience and/or model-estimated costs from FPInterface (FPInnovations) software, based on the supplemental knowledge provided by the LiDAR rasters. Included in this comparison, were cost savings associated with factors such as harvest block design, road construction, and differences in harvest volumes. This study, unique to Ontario, indicated that improved decision making based on the detailed spatial quantifications derived from LiDAR can result in substantial financial savings in planning and operational aspects of forest operations, to the extent that LiDAR costs may be recovered in less than two years. The study also suggested that such improved decision making may contribute to better stewardship of the landbase through a reduced environmental footprint of forest operations and the maximization value of the resource for the people of Ontario and for the industry involved.

References

Woods, M., Pitt, D., Lim, K., Nesbitt, D., Etheridge, D., Penner, M. and Treitz, P. 2011 "Operational implementation of a LiDAR inventory in Boreal Ontario" *Forestry Chronicle, 87(4):512-528*.

LiDAR and image based above-ground biomass estimation of non-forest trees: A comparative study

M. Rentsch[1], P. Krzystek[1] & A. Krismann[2]

[1] Munich University of Applied Sciences, Department of Geoinformatics, Karlstraße 6, 80333 Munich, Germany, rentsch@hm.edu, krzystek@hm.edu
[2] University of Hohenheim, Institute for Landscape- and Plant-Ecology, A.-v.-Hartmann-Straße 3, 70593 Stuttgart, Germany, a_krismann@uni-hohenheim.de

Paper Number: SL2012-034

Abstract

The main goal of the federal funded project 'LiDAR based biomass assessment' is the nationwide investigation of the biomass potential coming from landscaping activities of non-forest trees in Germany. Recent studies predict a great energetic potential of the harvested organic material which was more or less unused in the past. This paper is reporting about two methods to estimate the above-ground biomass (AGB) of non-forest trees from LiDAR data and aerial imagery.
First of all, mandatory field calibrations are performed for pre-defined grove types (e.g. 'hedge with trees', 'forest island'). For this purpose, selected reference groves are captured by airborne (ALS) and terrestrial laserscanning (TLS) in leaf-off and leaf-on conditions. The first method covers the determination of volume-to-biomass conversion factors (V2BCF) for each grove type which relate the reference AGB estimated from allometric functions with the vegetation volume using normalized digital surface models (nDSM). The second method is focused on a novel 3D Normalized Cut tree segmentation adopted for non-forest trees and the follow-on biomass calculation. The 3D segmentation provides for each tree segmented laser points from which a 3D tree model is derived using alpha shapes. A subsequent linear regression estimates the biomass using tree features calculated from the alpha shapes. The robustness of model is evaluated by a 5-fold cross-validation.
The methods are applied both to the LiDAR data (first/last pulse data (F+L) with 1 pt/m^2, full waveform data (FWF) with 10 pts/m^2) and to high-resolution aerial images (20 cm GSD). The investigations are applied on several reference groves of type 'hedge with trees' and 'forest island'. The V2BCF which is derived by the first nDSM-based method for five sample groves of type 'hedge with trees' results in a value of 0.0053 for FWF data and 0.0221 for F+L data. The V2BCF value for FWF corresponds well with the V2BCF of 0.0054 which was acquired from destructive sampling of six reference groves in Southern Germany (Baden-Württemberg). In this case, the reference (resp. the true) AGB could be determined exactly on-site and thus approves that the estimation of the reference AGB using allometric functions is appropriate. The ratio between the reference AGB and the segmentation-derived AGB by means of the second method leads to a value of 1.0430 for FWF data and 1.0590 for F+L data. The relative errors resulting from both methods indicate that the first method seems to be more suitable for high density LiDAR point clouds like FWF, and that the second method is superior in case of rather average point densities like for F+L data. Moreover, the different foliage conditions (FWF: leaf-on, F+L: leaf-off) have an impact on the results.
The second focus of our paper is on comparing the findings from LiDAR-based AGB estimates with the outcomes of dense stereo matching. For this purpose, the high-resolution aerial images - recorded in leaf-on conditions - are processed with a novel, commercial dense matching algorithm adopted from semi global matching.

G-LiHT:
Goddard's LiDAR, Hyperspectral and Thermal Airborne Imager

Bruce Cook[1], Lawrence Corp[2], Ross Nelson[3], Douglas Morton[4], Kenneth J. Ranson[5], Jeffery Masek[6] & Elizabeth Middleton[7]

[1]NASA Goddard Space Flight Center, bruce.cook@nasa.gov
[2]NASA Goddard Space Flight Center, lawrence.a.corp@nasa.gov
[3]NASA Goddard Space Flight Center, ross.f.nelson@nasa.gov
[4]NASA Goddard Space Flight Center, douglas.morton@nasa.gov
[5]NASA Goddard Space Flight Center, kenneth.j.ranson@nasa.gov
[6]NASA Goddard Space Flight Center, jeffrey.g.masek@nasa.gov
[7]NASA Goddard Space Flight Center, elizabeth.m.middleton@nasa.gov

Paper Number: SL2012-040

Abstract

Scientists at NASA's Goddard Space Flight Center have developed an ultra-portable, low-cost, multi-sensor remote sensing system for studying the form and function of terrestrial ecosystems. G-LiHT integrates two LIDARs, a 905 nm single beam profiler and 1550 nm scanner, with a narrowband (1.5 nm) VNIR imaging spectrometer and a broadband (8-14 µm) thermal imager. The small footprint (~12cm) LIDAR data and ~1 m ground resolution imagery are advantageous for high resolution applications such as the delineation of canopy crowns, characterization of canopy gaps, and the identification of sparse, low-stature vegetation, which is difficult to detect from space-based instruments and large-footprint LiDAR. The hyperspectral and thermal imagery can be used to characterize species composition, variations in biophysical variables (e.g., photosynthetic pigments), surface temperature, and responses to environmental stressors (e.g., heat, moisture loss). Additionally, the combination of LIDAR, optical, and thermal data from G-LiHT is being used to assess forest health by sensing differences in foliage density, photosynthetic pigments, and transpiration. Low operating costs (~$1 ha) have allowed us to evaluate seasonal differences in LiDAR, passive optical and thermal data, which provides insight into year-round observations from space. Canopy characteristics and tree allometry (e.g., crown height:width, canopy:ground reflectance) derived from G-LiHT data are being used to generate realistic scenes for radiative transfer models, which in turn are being used to improve instrument design and ensure continuity between LiDAR instruments. G-LiHT has been installed and tested in aircraft with fuselage viewports and in a custom wing-mounted pod that allows G-LiHT to be flown on any Cessna 206, a common aircraft in use throughout the world. G-LiHT is currently being used for forest biomass and growth estimation in the CONUS and Mexico in support of NASA's Carbon Monitoring System (CMS) and AMIGA-Carb (AMerican Icesat Glas Assessment of Carbon). For NASA's CMS, wall-to-wall G-LiHT data have been acquired over intensive study sites with historic LiDAR datasets, dense inventory data, stem maps and flux tower observations. For AMIGA-Carb, G-LiHT transects have been acquired over ICESat tracks and USDA-FS inventory plots throughout the CONUS, and similar data will be acquired in Mexico during 2013. This talk will highlight recent science results from continental-scale transects landscape-scale deployments of G-LiHT, as well as seasonal forest dynamics from repeat pass G-LiHT acquisitions.

Using Airborne Laser Scanning to constrain GORT Model Inversion

Henning Buddenbaum[1], Pyare Pueschel[1] & Joachim Hill[1]

[1]Environmental Remote Sensing & Geoinformatics, University of Trier, 54286 Trier, Germany. Email: buddenbaum@uni-trier.de

Paper Number: SL2012-048

Abstract

Biophysical properties on different scales have a large influence on the reflectance signal of forests. Clumping and shading effects from leaf level to stand level can lead to large differences in the reflectance of otherwise identical forests. These effects can most easily be studied using geometric-optical radiative transfer (GORT) models. Clumping and shading are also partly responsible for making the inversion of these models difficult, because there is no one-to-one-relationship between a small set of properties of interest and the reflectance signal. This ill-posed problem has been known for a long time but still remains to be solved. The larger the number of free parameters in model inversion, the more ambiguity arises in the results.
Conversely, the more is known about the variables that influence the reflectance signal or the more these variables can be reasonably constrained to narrow ranges, the better the estimation of the remaining variables via GORT model inversions.
Imaging spectroscopy (IS) alone, while providing access to species differentiation and state variables describing the ecophysiological conditions of forest stands, has limited capacities for deriving structural information. Airborne Laser Scanning (ALS), on the other hand, is not well suited for species detection or the derivation of biochemical parameters, but has the ability to measure structural parameters. Thus, for a detailed characterization of biophysical and biochemical forest parameters IS and ALS should be combined.
With ALS several structural forest variables can be estimated with acceptable to high accuracy; tree height, stem density, crown closure and leaf area index (LAI) being the most important ones that also influence the stand reflectance. Constraining these in the GORT model inversion process is expected to increase the estimation accuracy of remaining model parameters like chlorophyll and water content.
Our study site is in the Pfälzerwald forest near Kaiserslautern, Germany, at a core site of the forest research group within the EnMAP Core Science Team. In August 2009 two HyMap hyperspectral images of the area were acquired. The images were radiometrically and atmospherically corrected and geocoded. Additionally, an ALS data set of the area was acquired in September 2009. An ALTM-3100 sensor was used to cover the area with a mean point density of 4 m-2. The data was provided as ground, first, last, only and intermediate point clouds including intensity information. At the time of overflight forest structural measurements were carried out and branches were collected from the tree tops for analysis of biochemical leaf properties.
We inverted the GORT model INFORM, a combination of PROSPECT, SAIL, and FLIM. Spectra from the IS images were subjected to INFORM model inversion. Observation and illumination geometry, soil reflectance, skylight fraction and undergrowth LAI were set constant. We compare inversion results obtained by different levels of ALS assistance: (1) only IS data was used as input, no structural variables from ALS, (2) LAI was derived from ALS and fixed in the model inversion, and (3) all possible structural variables were derived from ALS and used in the model inversion.

A Dual-Wavelength Echidna® Lidar For Ground-Based Forest Scanning

Ewan Douglas[1], Jason Martel[2], Timothy Cook[3], Christopher Mendillo[4], Robert Marshall[5], Supriya Chakrabarti[6], Alan Strahler[7], Crystal Schaaf[8], Curtis Woodcock[9], Zhan Li[10], Darius Culvenor[11], David Jupp[12], Glenn Newnham[13], & Jenny Lovell[14]

[1]Boston University, Boston, MA, US, douglase@bu.edu
[2]University of Massachusetts Lowell, Lowell, MA, US, jfmartel@bu.edu
[3]University of Massachusetts Lowell, Lowell, MA US, Timothy_Cook@uml.edu
[4]Boston University, US, cmendill@bu.edu
[5]Stanford University, Palo Alto, CA, US, ram80@stanford.edu
[6]University of Massachusetts Lowell, Lowell, MA, US, supc@bu.edu
[7]Boston University, Boston, MA, US, alan@bu.edu
[8]University of Massachusetts Boston, Boston, MA, US, schaaf@bu.edu
[9]Boston University, Boston, MA, US, curtis@bu.edu
[10]Boston University, Boston, MA, US, zhanli86@bu.edu
[11]CSIRO Land and Water, Melbourne, VIC, AS, Darius.Culvenor@csiro.au
[12]CSIRO Marine and Atmospheric Research, Canberra, ACT, AS, david.jupp@csiro.au
[13]CSIRO Land and Water, Melbourne, VIC, AS, Glenn.Newnham@csiro.au
[14]CSIRO Marine and Atmospheric Research, Hobart, TAS, AS, Jenny.Lovell@csiro.au

Paper Number: SL2012-050

Abstract

A newly-constructed, ground-based lidar scanner designed for automated retrieval of forest structure, the Dual Wavelength Echidna Lidar (DWEL), separates laser "hits" of leaves from hits of trunks and branches using simultaneous laser pulses at 1548 nm, where leaf water content produces strong absorption, and at 1064 nm, where leaves and branches have similar reflectances. The DWEL uses a rotating mirror scan mechanism on a revolving mount, coupled with full digitization of return waveforms, to identify, locate, and parameterize scattering events in the three-dimensional space around the scanner.

In the DWEL instrument, the two measurement lasers are triggered simultaneously. Laser pulses are sharply peaked; they are expanded and collimated to a 6-mm beam diameter ($1/e^2$), then shaped into a top-hat cross section using a diffraction apparatus. Interchangeable optics provide a beam divergence of 1.25-, 2.5-, or 5-mrad. Scan resolution has three settings: 1-, 2-, and 4-mrad. Scan time varies with resolution: 11 min at 4 mrad; 41 min at 2 mrad; and 2.7 hr at 1 mrad. Two 0.5 mm InGaAs photodiodes measure the return signal, which is sampled by two digitizers at 2 gigasamples per second with 10-bit precision, providing 7.5-cm sampling of the 1.53 m pulse. Data are offloaded in real time via gigabit Ethernet to a separate field PC. Instrument weight is 20.4 kg (45 lbs), including high-density lithium ion batteries to meet the power requirement of 115 W. First images from the Harvard Forest, Petersham, MA, demonstrate the quality and new information content of DWEL scans.

The DWEL was built by the Center for Space Physics and Center for Remote Sensing, Boston University, with the support of NSF grant DBI-0923389, using a concept for a scanning lidar, trademarked Echidna®, developed and patented by CSIRO (US patent 7,187,452, Australian Patent 2002227768, New Zealand Patent 527547, Japanese Patent 4108478, and others).

Using LiDAR Data to Model Changes in Forest Structure from Mixed Severity Fires

Van R. Kane[1], James A. Lutz[2] & Malcolm North[3]

[1]School of Environmental and Forest Resources, University of Washington, vkane@uw.edu
[2]College of the Environment, University of Washington, jlutz@uw.edu
[3]U.S. Forest Service Pacific Southwest Research Station, mnorth@fs.fed.us

Paper Number: **#SL2012-056**

Abstract

Fire suppression over the last century has transformed many western U.S. conifer forests from open, patchy forests to more homogenous stands with continuous vertical and horizontal canopies. Managers often seek to reintroduce fire to these forests to reduce fuel loads and restore pre-suppression structure. However, they frequently lack information both on the present distribution of forest structures and how fire likely will change those structures. The high resolution and landscape coverage of airborne LiDAR could be an important tool for these managers, but frequently there are no corresponding field datasets with which to interpret the data. We have developed methods for using the LiDAR data itself to define forest structure classes that are ecologically meaningful and relevant to local conditions and goals of forest managers. We illustrate these methods through results from a recently completed study in Yosemite National Park. The Yosemite study area covered an area of 167 km^2 that experienced widespread mixed severity fires during our study period (1984-2010) but also retained substantial areas of unburned forest for comparison. We used a unique fusion of Landsat and airborne LiDAR data. Landsat measurements of the differenced Normalized Burn Ratio (dNBR) provided the history of fire severity over the 26 year period. From the LiDAR data, we measured the physical structure of forests in unburned patches and in patches that experienced different fire severities. Hierarchical cluster analysis was used to examine the LiDAR data and define five forest structural classes that differed in height, openness, and vertical fuel laddering. Random forest imputation was used to map these classes across the study area. We used the comparison of structure in unburned and burned patches to model how different fire severities changed structure across three forest types.

Using our structural metrics, we identified five statistically distinct vertical structure classes: open, spare, shorter, multistory, and top story. We also found three patterns of canopy patches and gaps associated with different fire severities: Canopy/gap patterns have gaps as small breaks in otherwise continuous canopy; Patch/gap patterns have similar proportions of interspersed canopy and gap; Open/patch patterns are open space with interspersed with small patches of canopy that are single trees or small tree clumps. Compared to stands outside fire perimeters, increasing fire severity resulted first in loss of canopy cover in lower height strata and increased number and size of gaps, then in loss of canopy cover in higher height strata with a canopy gap structure, and eventually to open areas with few or no trees. However, the fire severities at which these transitions occurred differed for each forest type. Our fusion of Landsat fire severity and LiDAR forest structure measurements allowed us to relate process (fire) to resulting change (forest structure). The results quantified the type and proportion of change that occurred with increasing levels of fire severity. LiDAR data collected over a large study area allowed us to clearly distinguish dominant patterns of change between forest types using LiDAR data as a primary measurement of forest structure.

Assessment of woody plant biomass with lidar and multispectral remote sensing

Nian-Wei Ku[1] & Sorin C. Popescu[2]

[1]Spatial Sciences Laboratory, Department of Ecosystem Science and Management, Texas A&M University, goofno17@neo.tamu.edu
[2]Spatial Sciences Laboratory, Department of Ecosystem Science and Management, Texas A&M University, s-popescu@neo.tamu.edu

Paper Number: #SL2012-061

Abstract

The encroachment of woody plants on grasslands causes the area of grasslands and croplands to shrink and reduce the yield of forage crops and agriculture products. Nevertheless, the woody plants provide avian and small mammal habitats and enriche biodiversity on grasslands. In Texas, honey mesquite (*Prosopis glandulosa* Torr.) is one of the main species of woody plants in the woodland ecosystem. The objective of this study are 1) to use airborne lidar and multispectral remote sensing data to assess regional scale honey mesquite biomass, 2) to evaluate the differences of the projected available biomass between those two remote sensing data. The normalized difference vegetation index (NDVI) was derived from high spatial resolution multispectral remote sensing data and the ground reference data were used with cokriging interpolation methods to estimate local and regional scale honey mesquite biomass in a rangeland area and create a biomass map. Additionally, the height of vegetation has been found to have a strong relationship with the vegetation biomass. The height bin data was acquired from the lidar point cloud data. The point density of each bin represented the quantity of biomass. Thus, airborne lidar data of the same area was used with the ground reference data to estimate biomass and create a biomass map.

Combining LiDAR and hyperspectral remote sensing data to improve information extraction for forestry

P. Kempeneers[1], F. Devriendt[2], F. Van Coillie[2] K. Vandekerckhove[3] & F. Morsdorf[4]

[1]VITO, TAP pieter.kempeneers@vito.be
[2]Universiteit Gent Flore.Devriendt@Ugent.be
[2]Universiteit Gent Frieke.VanCoillie@UGent.be
[3]INBO kris.vandekerkhove@inbo.be
[4]Remote Sensing Laboratories felix.morsdorf@geo.uzh.ch

Paper Number: SL2012-063

Abstract

The objective of this study was to investigate the potential of the combination of LiDAR and hyperspectral data to extract reliable forest parameters, such as tree species composition, stand diversity and forest vitality. This presentation will focus on important processing steps to obtain this information and will present the results for the different forest sites.

LiDAR and hyperspectral remote sensing data were acquired over three different forest sites in Belgium. LiDAR data was obtained from a TopoSys sensor Harrier 56 at full waveform, with a point density of three to four points per square meter. Two hyperspectral data sets were available, based on different sensors. The first hyperspectral sensor was the Compact Airborne Spectographic Imager (CASI), covering the visual and near infrared part of the spectrum (368 - 1551 nm). The second was the newly developed imaging spectrometer by a Swiss-Belgian consortium on behalf of ESA, the Airborne Prism Experiment (APEX). It covers a wider spectral range, including the short wave infrared (372-2498 nm) with more spectral bands. On the other hand, spatial resolution was better for CASI (1 m) than for APEX (2 m).

The processing steps involved the hyperspectral data pre-processing (radiometric and geometric correction) and extraction of forest structure information from the LiDAR data. The fusion of hyperspectral and LiDAR data is relatively new and most studies have focused on (e.g., tree species) mapping. Here we show the potential of combining both sources of data for 1. the pre-processing of hyperspectral data and 2. the extraction of forest structure information. As an example, LiDAR data helped to improve the geometrical correction of hyperspectral imagery. LiDAR data were also involved in the training process of the supervised classification of tree species. Pure reflectance spectra were extracted from sunlit tree crowns, which were delineated from a vegetation canopy height model based on the LiDAR data. To show the potential of the combination of hyperspectral and LiDAR data, results will be presented for the forest parameters obtained with and without data fusion.

VoxLAD: a model for rapid estimation of leaf area distributions in broadleaved trees from terrestrial LiDAR measurements

Beland, M.[1,3*], Widlowski, J.-L.[2], Fournier, R.A.[1], Baldocchi, D.D.[3], Verstraete, M.M.[2]

[1] Depart. of Applied Geomatics, University of Sherbrooke, Canada
[2] Institute for Environment and Sustainability, Joint Research Center (JRC), European Commission, Ispra, Italy
[3] Depart. of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA
[*] Corresponding author: mbeland@berkeley.edu

Paper number: #SL2012-076

Abstract

Tree leaf area refers to a key vegetation structure parameter for the study of ecosystem functioning and land-atmosphere interactions. Terrestrial LiDAR Scanners (TLS) hold great potential in recording spatially explicit leaf area density (LAD) with a high level of detail. Studies on this topic have highlighted several factors affecting leaf area density estimates. These factors are (1) the presence of non-photosynthetic material in point clouds, (2) leaf inclination angle, (3) occlusion effects (gaps in data caused by areas being shaded by plant parts), (4) point density changing with distance from scanner, and (5) laser beam size. A method addressing the above points has been presented in Beland *et al*. (2011). Recent developments in faster TLS instruments has reduced constrains related to acquisition time; this opens the door to an increase in the number of measurements from different scanning positions, either to extend the coverage of the scanned area or reduce occlusion effects in dense canopies. In order for models to handle the amount of data generated, we need fast models able to minimise the number of computations, and ideally process the data using parallel computing. Here we present the VoxLAD model, which is based on the same voxel matrix framework as the model of Beland *et al*. *(*2011), but relies on computational geometry instead of ray tracing computations, and is capable of using multiple computer cores for processing and managing memory. The algorithm developed relates the number of points (or laser returns) in a voxel to the number of laser pulses entering the voxel. The Beer-Lambert law is then used to derive LAD based on laser pulse transmission through the voxel. An experiment conducted to assess the level of clumping at the voxel scale –typically 30 cm in side length- is also presented. The model's accuracy was tested against direct field measurements, and results show that it can be used to accurately estimate the 3D LAD distribution within canopies. This model represents great potential for usage within studies looking at the calibration and validation of products derived from aerial LiDAR measurements as well as other optical earth observing systems. The model is coded in Matlab and can be made available upon request addressed to the corresponding author.

References

Béland, M., Widlowski, J.-L., Fournier, R., Côté, J.-F. and Verstraete, M.M., 2011. Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements. Agric. For. Meteorol., 151(9): 1252-1266.

Classification of tree types using countrywide airborne laser scanner data, color infrared images and field data

Johannes Schumacher[1], Sorin C. Popescu[2] & Thomas Nord-Larsen[3]

[1]University of Copenhagen, Forest & Landscape, josc@life.ku.dk
[2]Texas A&M University, Spatial Sciences, s-popescu@tamu.edu
[3]University of Copenhagen, Forest & Landscape, tnl@life.ku.dk

Paper Number: SL2012-083

Abstract

The demand for efficient collection of information about forest resources and its allocation has been the driving force for research being conducted on remote sensing techniques in forest inventory in the last years. Remote sensing data are usually analyzed on plot level (area based) or on single tree level (object based). In contrast to the object based approach the area based approach has reached operational status in some countries. However, for an accurate forest resource estimation knowledge of tree types and their composition is essential but cannot be derived from data collected on plot level.

We present an approach for an object based classification of forest inventory plots into broadleaf, coniferous and mixed ones using remote sensing data. The data set consists of countrywide airborne laser scanner (ALS) data (0.5 points/m^2), countrywide color infrared (CIR) images (near infrared, red and green band), national forest inventory (NFI) data and field data from a species trial in Denmark. The ALS data were collected during leaf-off season in 2006/07. The NFI data and the data from the species trial were collected in the same period. The CIR images were taken during leaf-on season between middle of May and middle of July 2010.

Model development is conducted on the plots of the species trial where the species of each tree is precisely known. The model is applied on NFI plots. In the first step a segmentation of single trees or tree groups of the same type is conducted using segmentation tools of the Trimble eCognition software. In the second step the classification of each object is performed. The segmentation is based on the spectral information of the CIR images and the raw ALS points. Each segment is separated into sub-objects of sunlit and shadowed part. For each whole segment ALS based variables and for each sunlit part of each segment CIR based variables are derived. ALS based variables include the proportion of last pulse hits on vegetation and on the ground and height percentiles and density metrics of all pulses. CIR based variables include minimum, mean and maximum band values and HSI (hue, intensity, saturation) transformation values of the multispectral data. An unsupervised fuzzy classification of each segment into broadleaf or conifer is performed. Finally each plot is assigned to one of three categories broadleaf, conifer or mixed depending on the fraction of area covered by broadleaf and conifer segments.

Preliminary classification results on the species trial based on a limited set of variables showed an overall accuracy of 80% and a kappa coefficient of 0.5. This pre-stratification into broadleaf, conifer or mixed plots allows for the selection of an appropriate model to estimate forest resources of pure or mixed stands.

Small area estimation using ALS data

Johannes Breidenbach[1], Edgar Kublin[2] & Rasmus Astrup[1]

[1]Norwegian Forest and Landscape Institute, job@skogoglandskap.no
[2]Forest Research Institute of Baden-Württemberg

Paper Number: #SL2012-086

Abstract

The estimation of forest attributes (e.g., timber volume or biomass) for small units such as stands is a traditional field of airborne laser scanning (ALS) based forest inventories. Such ALS-based forest inventories (AFIs) were among the first operational applications of ALS within the forestry realm. AFIs rely on the availability of wall-to-wall ALS data and field sample plots that are used to model the relation between field observations and ALS metrics.

In general, small area estimation (SAE) provides a large set of model-dependent tools to estimate forest parameters in AFIs (Rao 2003). In practise, however, the synthetic estimator appears to be used most frequently. In addition, uncertainties arising from statistical models and the variability of the ALS data in the stand are frequently neglected. This can lead to acceptance problems and biased projections.

In our study area, Lardal in south-eastern Norway, 30 stands were randomly selected. Trees on seven sample plots with 8.92 m radius were measured within each stand according to the protocols of the Norwegian National Forest Inventory (NNFI). In addition, 31 NNFI sample plots on a 3x3 km grid were available. High-density ALS data were used as auxiliary information. A linear mixed-effects model (LMEM) with a random intercept on the stand level was fit to predict biomass per hectare. The residuals based on the fixed effects only had a clear correlation on stand level which would result in biases of synthetic estimates. It should be noted that this bias is not a result of the clustered sampling and similar biases are expected if all sample plots were systematically distributed over the study area.

The use of a LMEM is the base for the empirical best linear unbiased prediction (EBLUP) estimator which combines field- and remote sensing data in an optimal manner (Breidenbach & Astrup 2012). The measurement of few sample plots within a stand was sufficient to calibrate the random effect which reduces the bias dramatically. In practise, this can be especially interesting for high-valued stands that are close to the final harvest. Hence, the EBLUP estimator provides a good possibility for combining Forest Management Inventories and National Forest Inventories allowing synergistic use of resources (field plots and remotely-sensed data) on both levels.

References

Breidenbach, J. & Astrup, R., 2012. Small area estimation of forest attributes in the Norwegian National Forest Inventory. *European Journal of Forest Research*, 131, 1255-1267.

Rao, J.N.K., 2003. Small area estimation, Wiley-Interscience.

TLS-inventory and variable radius sample design for orest inventory in Germany

Andreas Fritz[1] Holger Weinacker[1] & Barbara Koch[1]

[1]FeLis, Albert-Ludwigs University Freiburg, Germany,
andreas.fritz@felis.uni-freiburg.de

Paper Number: SL2012-088

Abstract

Terrestrial Laser Scanners are becoming lighter, faster and cheaper and are therefore becoming increasingly popular in forest inventory concepts. Many recent studies have focussed on stem detection under more or less controlled conditions. This includes the registration of multiple scans or scans optimally positioned within a stand to guarantee good detection results. In this study we are evaluating the implementation of TLS-systems in the national forest inventory of Germany for stem sample size and stem attributes. Special focus lies on the effect of occlusion. National inventories rely on grid based permanent sample plots where repetitive surveys are conducted. In Germany, the concept of variable radius sample design is applied. In this design, sample trees are chosen based on their distance to the plot centre and their diameter at breast height (dbh). To ensure result comparability and be consistent with national inventory design, the scanner should be placed on plot centres. These are not necessarily good positions for TLS-scans, for example in the case of an old tree close to the plot centre resulting in high occlusion. Single scan mode was chosen for this study because we believe this less time-demanding method is most realistic for such inventories. We used a cluster algorithm to automatically determine trees in the point cloud and derive base inventory parameters (dbh, height). Before clustering, ground point classification and filtering were applied to the point cloud. The clustering approach used is based on neighbourhood distances in 3D-space and a distance threshold. The shape of the clusters is then inspected by testing their points against cylinder geometry. Clusters meeting the geometrical requirements are then classified as trees and further processed to derive inventory parameters. The results have been compared to official inventory results mapped in the same time period as the TLS-acquisition (Winter 2011/2012). 14 sample plots were analysed in stands dominated by deciduous trees. The plots differ in terms of understory and age. National inventory mapped between 1...11 sample trees per plot. We found between 0…10 sample trees over all in the TLS-scans. In total, 83% of the detected trees were the same as those chosen by the variable radius sample design of the manual inventory. When comparing the attributes of the manual inventory with the data derived from our automatic tree detection algorithm we can conclude that TLS-measurements on fixed position plots are a suitable tool to include in the standard inventory procedure. The occlusion effect seems to be not much influencing the tree selection. In addition, TLS-data contains additional information to that collected by traditional inventories, such as taper, sweep and branchiness.

Quantification of biomass change in young forest using airborne laser scanner data

Roar Økseter[1], Ole Martin Bollandsås[1], Terje Gobakken[1] & Erik Næsset[1]

[1]Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432 ÅS, Norway; E-mail: roar.okseter@umb.no

Paper Number: 092

Abstract

In order to comply with the reporting requirements of the UN Climate Convention in countries where data from field surveys are lacking like for example in many tropical countries, there is a need for an efficient tool to monitor changes in biomass. To obtain data in a cost-efficient way over large areas, use of remote sensing techniques is required. Airborne laser scanning (ALS) is a promising remote sensing technique to assist estimation of biophysical properties of trees and forests. Previous studies using ALS-data for forest inventory purposes has mainly focused on mature forests because timber has been the most important resource. However, because biomass and biomass changes are important related to changes in carbon stocks, forests of all ages will be of interest. The aim of this study was to explore the capability of multi-temporal airborne laser scanner to quantify the change of total aboveground tree biomass in young forests. The precision and accuracy of two different methods were compared.

The study area is located in the municipality of Våler, southeastern Norway. The field data consists of 39 circular plots from young forest, each 200 m^2 in size. Field and ALS data were collected in 1999 and re-measured in 2010. The data cover 11 growth seasons, and the mean change of biomass during this period was 49.14 Mg/ha with an initial mean biomass of 6.63 Mg/ha.

Total aboveground biomass and biomass change were modeled using explanatory variables derived from ALS data with ordinary least squares. Different thresholds to separate between vegetation and ground echoes were tested, but the results revealed no major differences between the choices of thresholds. However, a low threshold will capture more of the variability in young forests, and a threshold of 0.5 m was therefore chosen. Only the first return of the ALS data was considered in this study.

The first method (Method 1) estimates change in biomass by means of two models, i.e. one model for predicting biomass at each point in time. The change was estimated as the difference between the predicted biomass in 2010 and predicted biomass in 1999. The second method (Method 2) was a direct estimation of biomass change. The observed differences between the first return height percentiles and density variables obtained from ALS data on the two points in time were used as explanatory variables and the observed biomass change was used as the response.

The two methods were evaluated using cross validation. The differences between the estimated and the observed biomass change, for both methods, were not significantly different from 0. The RMSE values were around 20 % of the mean observed biomass change for both methods. Thus the two methods seemed to be quite similar in terms of accuracy, but Method 1 appeared to be slightly better than Method 2. The choice of method for quantification of biomass change will probably be dependent on the type of forest.

Forest inventory using LiDAR at Bergvik Skog AB

Lars Sängstuvall[1], Ingemar Gillgren[1], Ola Lindgren[2], Erik Heimdal Iversen[3], Thomas Brethvad[3]

[1] Bergvik Skog AB, lars.sangstuvall@bergvikskog.se
[2] OL Skogsinventering AB
[3] COWI

Paper Number: SL2012-093

Bergvik Skog AB owns 1.9 million hectares of productive forest land in Sweden, corresponding to 8% of all Swedish productive forest land. The annual felling amounts to more than 6 million m3. Stand-level forest information is stored in a GIS-based stand register comprising 245,000 stands. The stand register information is used for various forest planning activities at the strategical, tactical and operational levels, representing different time horizons and complexity of decision problems. The information has been acquired over 20-30 years and using different field inventory and remote sensing methods, and hence the accuracy of the information is varying, but generally poor. For example, the standard error for stand-level standing volume per hectare in the stand register is 25-40%. Until recently, field-based sample-plot inventories were done following all forest management activities, to update the stand register with information to serve as a basis for following planning activities. The inventory cost of such sample-plot inventories ranges from $5 to $30 per hectare, and the expected standard error for e.g. standing volume varies around 10-15%.

Estimating forest variables from LiDAR data using the area-based method renders similar or higher accuracy than that of sample-plot inventories for forest variables such as standing volume and mean height. When also considering the fairly low price for such data, ranging from $2 to $4 per hectare, LiDAR-based forest inventories stand out as very competitive compared to sample-plot inventories. Hence Bergvik Skog AB has decided to describe all of its forest land using the area-based method, and this presentation will briefly summarize how we incorporate and utilize the new data in our planning activities.

The new datasets obtained from the LiDAR inventory includes a vector data 15×15 m square grid dataset containing forest variables, an automatically updated stand delineation based on a vegetation height raster, a new digital terrain model and some derived terrain layers. The new stand delineation and stand mean variables usually replaces the old information in the stand register. However, stand information which has been acquired and updated post-scanning (due to e.g. harvesting activities) is generally preserved. This poses a challenge, since we want to replace novel but poor as far as possible, to avoid discarding high-quality LiDAR-based forest data which in fact are more accurate, although older.

This updated stand register information serves as a base for improved forest- and stand-level strategic and tactical planning. For tactical and operational aspects – such as identifying dense areas for thinning, regardless of stand delineation – processed versions of the grid are used in new work routines. Replacing old, poor stand-level information with high-accuracy, spatially comprehensive information poses a great challenge both technically and organizationally. To make the most out of our investment in new information, and to identify and fully realize its potentials, much more work still remains.

Operational Inventory from low density airborne LiDAR: approaches for model estimation

M.Penner[1], M. Woods[2] & D.G. Pitt[3]
[1]Forest Analysis Ltd. mpenner@vianet.ca
[2]Ontario Ministry of Natural Resources murray.woods@Ontario.ca
[3]Canadian Wood Fibre Centre dpitt@NRCan.gc.ca

Paper Number: SL2012-104

Operational Inventory from low density airborne LiDAR: approaches for model estimation

Non parametric techniques (kNN and random forest) are used to produce a 400-m^2 raster-based forest inventory from relatively low densities of LiDAR (~ 1 pulse per m^2) and the results contrasted with traditional parametric regression. Previous work with traditional parametric regression demonstrated significant differences in prediction models by forest type (Woods et al. 2011), necessitating either prediction of forest type from LiDAR or another source such as manual interpretation from aerial photographs. Random Forest and kNN are common non-parametric techniques and are less reliant on prior stratification. Random Forest, in particular, is shown to perform well compared to parametric regression and kNN. The advantages and disadvantages of random forest, kNN imputation and parametric regression are contrasted on the Hearst Forest in boreal Ontario, Canada.

References

Woods, M., Pitt, D, Penner, M, Lim, K., Nesbitt, D., Etheridge, D., and Treitz, P. 2011. Operational implementation of a LiDAR inventory in Boreal Ontario. For. Chron. 87(4):512- 528.

Branch growth and tree change detection using terrestrial laser scanning

Krooks, Anssi[1]; Kaasalainen, Sanna[1]; Raumonen; Pasi[2], Kaasalainen, Mikko[2]; Kaartinen, Harri[1]; Kukko, Antero[1]; Puttonen, Eetu[1]; Liski, Jari[3]; Holopainen, Markus[4]; Vastaranta, Mikko[4]

[1]Department of remote sensing and photogrammetry, Finnish Geodetic Institute, Geodeetinrinne2, P.O. Box 15, 02431 Masala, Finland. anssi.krooks@fgi.fi
[2]Dept. of Mathematics, Tampere University of Technology, P.O. Box 553, FI-33101 Tampere, Finland
[3]Finnish Environment Institute, Mechelininkatu 34a, P.O. Box 140, FI-00251 Helsinki, Finland
[4]Department of Forest Sciences, University of Helsinki, P.O.Box 27, FI-00014 Helsinki, Finland

Paper Number: #SL2012-105

Abstract

We have developed a computational method to produce automatically comprehensive and quantitative models of trees from terrestrial laser scanning derived point clouds (Raumonen et al., 2011). Tree surface is modeled locally with cylinders, whose size, orientation, and location describe the local geometric details of the tree. Also the 3D branching structure is contained in the model. From the model we can easily calculate, e.g., over-the-ground volume and branch size distribution. Also more detailed information such as diameter profiles of single branches and branching angles are easily available from the model. These tree characteristics can be used, e.g., for models of carbon footprint from decaying vs. energy use.

This paper describes an application of TLS data and the tree modeling to produce 4D tree models by time series of tree growth, change, and littering during growing seasons in quantitative level. Many of these characteristics have been difficult or even impossible to measure operationally: the measurement requires cutting (which then limits the study to a single measurement instead of, e.g., time series) and laborious manual and field measurements, some of them very hard to perform efficiently or accurately.

One way to estimate tree parameters, such as biomass, is to use statistical models developed from detailed manual measurements. The simplified statistical models are limited in accuracy, and cannot give accurate quantitative information on a single tree, and certainly not for a single branch or other more detailed regions of the tree. We show that our method, based on TLS and accurate local cylinder approximation of three branch structure, is capable of producing detailed information on the branch-level changes in trees, such as growth and littering over time. In this paper, we present the first results and a validation scheme using TLS and manually measured changes in branch volume.

References

Raumonen, P., Kaasalainen, S., Kaasalainen, M. and Kaartinen, H., 2011. Approximation of volume and branch size distribution of trees from laser scanner data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Calgary, Canada,* Vol 38 (5/W12)

Direct Estimation of Leaf Area Index and Foliage Area Volume Density Profile through Voxelization of Point-Cloud Forest Reconstructions from Ground-Based Full-Waveform Echidna® Lidar Scans

Xiaoyuan Yang[1], Alan H. Strahler[2], Crystal B. Schaaf[3] , Zhan Li[4], Tian Yao[5], Feng Zhao[6], Zhuosen Wang[7], Curtis Woodcock[8], David Jupp[9], Darius Culvenor[10], Glenn Newnham[11], Jenny Lovell[12] & Wenge Ni-Meister[13]

[1] University of Massachusetts Boston, Boston, MA USA & Boston University, Boston, MA USA and Xiaoyuan.Yang@umb.edu
[2] Boston University, Boston, MA USA and Alan@bu.edu
[3] University of Massachusetts Boston, Boston, MA USA & Boston University, Boston, MA USA and Crystal.Schaaf@umb.edu
[4] Boston University, Boston, MA USA and zhanli86@bu.edu
[5] Montclair State University, Montclair, NJ USA and yaot@mail.montclair.edu
[6] University of Maryland, College Park, MD USA and fengbjfu@gmail.com
[7] University of Massachusetts Boston, Boston, MA USA & Boston University, Boston, MA USA and Zhuosen.Wang@umb.edu
[8] Boston University, Boston, MA USA and Curtis@bu.edu
[9] CSIRO Marine and Atmospheric Research, Canberra, ACT, Australia and David.Jupp@csiro.au
[10] CSIRO Land and Water, Melbourne, VIC, Australia and Darius.Culvenor@csiro.au
[11] CSIRO Land and Water, Melbourne, VIC, Australia and Glenn.Newnham@csiro.au
[12] CSIRO Marine and Atmospheric Research, Hobart, TAS, Australia and Jenny.Lovell@csiro.au
[13] Hunter College of the City University of New York, New York, NY, USA and wnimeist@hunter.cuny.edu

Paper Number: SL2012-107

Abstract

A new method for retrieval of leaf area index (LAI; m^2) and foliage area volume density (FAVD; m^2 LAI/m^3 volume) of forest stands uses voxels derived from a three-dimensional point cloud of scattering events observed in registered forest scans using a full-waveform, ground-based lidar, the Echidna® Validation Instrument (EVI). The voxelization process uses the attributes of scattering events, including the gap probability to the event, its apparent reflectance, and the volume associated with the return of the laser pulse, to produce the volumetric dataset. Classification procedures(Yang *et al.* 2012), based on the shape of the laser pulse returned to the instrument, separate trunk from foliage scattering events. Measurements or estimates of leaf reflectance, leaf-to-fine-branch ratio, and clumping at the scale of the pulse width or finer are required. Leaf angle distribution is accommodated with a simple model based on gap probability with zenith angle as observed in individual scans of the stand. An important advantage of the method is that coarse-scale (between-crown) clumping is observed directly and does not require parametric correction. For validation, we compare LAI and FAVD profiles retrieved directly from the voxelized 3-D forest reconstructions with those observed in airborne and field measurements.

Previous studies on single EVI scans using azimuth-averaged gap probability with zenith

angle (Jupp *et al.* 2009; Zhao *et al.* 2011) have shown LAI and FAVD estimations to agree well with those of traditional hemispherical photos and LAI-2000 measurements. By merging point clouds constructed from overlapping EVI scans, vegetation stands are reconstructed in 3-D space (Strahler *et al.* 2008; Yang *et al.* 2012), allowing virtual direct representation of biomass distribution.

Voxelized 3-D forest reconstructions derived from EVI point clouds provide a pathway to estimate "ground truth" FAVD, LAI, and biomass without destructive sampling. They can be used to validate large-footprint spaceborne and airborne lidar systems, thus facilitating large-area inventories. The enhanced characterization of leaf area distribution is of interest to both land biogeoscientists who require bulk vegetation biomass measures and to atmospheric biogeoscientists, who require information on surface roughness, photosynthesis, and respiration processes. Moreover, the EVI can be deployed to monitor disturbance and deforestation detected by optical sensors, such as MODIS or Landsat, to provide better calibration of the type and nature of change.

References

Jupp, D.L.B., Culvenor, D.S., Lovell, J.L., Newnham, G.J., Strahler, A.H., & Woodcock, C.E. (2009). Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna®'. *Tree Pysiology, 29*, 171-181

Strahler, A.H., Jupp, D.L.B., Woodcock, C.E., Schaaf, C.B., Yao, T., Zhao, F., Yang, X., Lovell, J.L., Culvenor, D.S., Newnham, G.J., Ni-meister, W., & Boykin-morris, W. (2008). Retrieval of forest structural parameters using a ground-based lidar instrument (Echidna®). *Canadian Journal of Remote Sensing, 34,*, 5426-5440

Yang, X., Strahler, A.H., Schaaf, C.B., Jupp, D.L.B., Yao, T., Zhao, F., Wang, Z., Culvenor, D.S., Newnham, G.J., Lovell, J.L., Dubayah, R.O., woodcock, C.E., & Ni-Meister, W. (2012). Three-Dimensional Forest Reconstruction and Structural Parameter Retrievals Using a Ground-Based Full-Waveform Lidar Instrument (Echidna®). *Remote Sensing of Environment, submitted.*

Zhao, F., Yang, X., Schull, M.A., Roman, M., Tian, Y., Wang, Z., Zhang, Q., Jupp, D.L.B., Culvenor, D.S., Newnham, G.J., Ni-Meister, W., Schaaf, C.B., Woodcock, C.E., Strahler, A.H., & Richardson, A. (2011). Measuring effective leaf area index, foliage profile, and stand height in New England forest stands using Echidna® Validation Instrument (EVI) ground-based lidar. *Remote Sensing of Environment, 115*, 2954–2964

Three-dimensional Reconstruction of *Eptesicus fuscus* Flight Trajectories Combining Thermal Imaging and Ground-based Lidar

Xiaoyuan Yang[1], Crystal Schaaf[2], Alan Strahler[3], Thomas Kunz[4], Nathan Fuller[5], Margrit Betke[6], Zheng Wu[7], Zhuosen Wang[8], Diane H. Theriaul[9], Darius Culvenor[10], David Jupp[11], Glenn Newnham[12], Jenny Lovell[13].

[1] University of Massachusetts Boston, Boston, MA USA & Boston University, Boston, MA USA and Xiaoyuan.Yang@umb.edu
[2] University of Massachusetts Boston, Boston, MA USA & Boston University, Boston, MA USA and Crystal.Schaaf@umb.edu
[3] Boston University, Boston, MA USA and Alan@bu.edu
[4] Boston University, Boston, MA USA and kunz@bu.edu
[5] Boston University, Boston, MA USA and nwfuller@gmail.com
[6] Boston University, Boston, MA USA and betke@bu.edu
[7] Boston University, Boston, MA USA and wuzheng@bu.edu
[8] University of Massachusetts Boston, Boston, MA USA & Boston University, Boston, MA USA and Zhuosen.Wang@umb.edu
[9] Boston University, Boston, MA USA and deht@cs.bu.edu
[10] CSIRO Land and Water, Melbourne, VIC, Australia and Darius.Culvenor@csiro.au
[11] CSIRO Marine and Atmospheric Research, Canberra, ACT, Australia and David.Jupp@csiro.au
[12] CSIRO Land and Water, Melbourne, VIC, Australia and Glenn.Newnham@csiro.au
[13] CSIRO Marine and Atmospheric Research, Hobart, TAS, Australia and Jenny.Lovell@csiro.au

Paper Number: SL2012-108

Abstract

The nature of forest structure (e.g., canopy cover, distribution of overstory and understory, foliage density, forest gaps, stand-scale variability, forest fragmentation, availability of standing deadwood, and forestry practices) is an important contributor to the maintenance of successful bat populations. In order to study bat foraging behavior with respect to forest structure, we combined thermal imaging technology with a ground-based lidar system to reconstruct *Eptesicus fuscus* (big brown bats) flight trajectories in three-dimensional space.

Flight trajectories of bats were reconstructed using imaging data from FLIR ThermoVision® SC8000 cameras at a bat roosting and maternity site in Petersham, MA, and co-registered to a three-dimensional forest reconstruction built from nine Echidna® Validation Instrument (EVI) scans at the site. Patterns of flight trajectories during first five seconds of emergence from the roosting barn showed the bats chose different flight routes to forage along the edge of a forest and into the understory.

Study of the flight velocity and height categorized the 24 identified flight trajectories of bats into different behavior groups. Results of the five second period immediately after they emerged from the barn showed that during the first second, bats speeded up and dropped height. In the 2nd second, they slowed down and adjusted flight direction. In the third second, some bats slowed down further while others speeded up as they all continued to drop in height. They spread into forest following the fourth second of flight. Although all bats were guided by echolocation, different groups of bats were traced along different flying routes to avoid hitting a tree that was in their flight path.

This study is an initial demonstration that breaks new ground for future ecological studies in which flight trajectories of bats are coupled with the canopy reconstructions to better establish their responses to different habitat characteristics.

The Carbon Monitoring System: preliminary results of biomass inventory using small footprint LiDAR

Juan Suárez[1,2,3], Jacqueline Rosette[2,3], Naiara Pinto[2], Ralph Dubayah[2], Bruce Cook[3] and Temilola Fatoyinbo[3]

[1]Forest Research, Northern Research Station, Roslin, Midlothian EH25 9SY, United Kingdom, juan.suarez@forestry.gsi.gov.uk

[2]Geography Department, University of Maryland, College Park, MD, 20742, USA; npinto@umd.edu; dubayah@umd.edu

[3]NASA Goddard Space Flight Center, Biospheric Sciences, Greenbelt, MD, 20771, USA. ross.f.nelson@nasa.gov; jacqueline.rosette@nasa.gov; temilola.e.fatoyinbo@nasa.gov; bruce.cook@nasa.gov

Paper Number: ### SL2012-110

Abstract

The Carbon Monitoring System (CMS) project started in September 2010 with a mandate from the U.S. Congress to NASA. The primary goal of the project is the estimation of above-ground biomass stocks and carbon fluxes, the appraisal of error budgets and the delivery of cartographic products and other geo-referenced information for selected locations at a scale that is ecologically meaningful and relevant for forest management. As part of an effort to develop a methodology of analysis universally applicable across the USA, small footprint LiDAR data have been analysed for two counties in Maryland and a plantation forest in North Carolina.

Field data in Maryland consisted of 301 plots of variable size determined by different Basal Area Factors (BAF) according to vegetation structure (BAF-10 for young growth, BAF-20 for grown areas with understorey presence and BAF-30 for grown areas and no understorey). Plot selection was undertaken after a stratification of the Canopy Height Model (CHM), derived from airborne LiDAR (taken in 2004) and divided into 3 height classes (0-5, 5-25 and more than 25 m), and their interception with the National Land Cover Database 2006 (NLCD) classes: conifers, broadleaves, mixed, wetlands and non-Forest. Field data in Parker Track consisted of just one stratum (Conifers) and 35 mensuration plots of 7.31 m radius randomly distributed. LiDAR data were taken in 2011 with NASA's new G-LiHT system. Biomass in the field was calculated as a function of DBH and species following the method described by Jenkins et al., 2003.

An area-based analysis was performed in these two areas using the extent of the NLCD pixels (30x30 m^2). For each cell, LiDAR metrics were calculated containing: percentile distribution of heights (bin 5%), summary statistics, percentage of canopy cover above 2.5 m and density values (bin 5%). A model fit was obtained by a sigmoidal function (Boltzmann fit) using H_{60} (60th percentile) corrected by Percentage of Canopy Cover. This approach intended to couple LiDAR analysis with more biological foundations as described in Bertalanffy's growth model (Pienaar and Turnbull, 1973). As the Boltzmann model reduces complexity to four parameters, this system can be implemented nationwide to monitor carbon. The A-parameter was set to the maximum height

percentile value (i.e. maximum biomass value in each stratum) and the B-parameter was fixed at the lowest height percentile value (i.e. 0 or no biomass). Therefore, the model calculated only the slope (D) and the 50th (median) percentile (C) parameters. In the context of a national Carbon Monitoring System, D can be associated to changes in forest structure and C to sources and sinks of carbon when working with time-series observations.

The results showed coefficients of determination ranging above 0.7 for all the strata with the exception of mixed woodlands (c. 0.5). The distribution of the residuals showed a bias for plots with largest biomass levels measured in the field, due to the unintended inclusion of boundary trees. The final results were converted into a 30m raster map showing the spatial distribution of biomass in Mg ha^{-1}.

Using terrestrial lidar to quantify gap fraction in boreal forests

Steven Hancock[1], Tim Reid[2], Richard Essery[2], Joel Carle[3], Robert Baxter[4], Brian Huntley[4]

[1]Swansea University, Geography and National Centre for Earth Observation
s.hancock@swan.ac.uk
[2]University of Edinburgh, Geosciences
[3]Geospatial Research ltd, Durham
[4]Durham University, Biological and Biomedical sciences

Paper Number: #SL2012-129

Abstract

The work presented here is part of a larger project which aims to improve numerical models used for weather, climate and water availability predictions. Land surface models (LSMs) are used in general circulation models to make predictions of climate and water availability. Snow has a dramatic effect upon these processes, however it is a known weakness of LSMS, particularly in forests (Rutter et al 2009). Part of this weakness comes from predicting snow melt as incoming radiation, air temperature and land cover change. Radiative transfer schemes allow us to determine how forests interact with radiation (both long and shortwave) and how much reaches the snow and contributes to the melt energy.

A terrestrial lidar is capable of measuring the full structure of a forest canopy in far more detail than any other practical method. This allows the creation of a complex radiative transfer (RT) model which can be used to test more efficient models for use in an LSM. The part of the project presented here concentrates on the creation and validation of the complex RT model.

A terrestrial laser scanner (Leica C10 first return system) was taken to Abisko in Northern Sweden (sparse birch forests) and Sodankyla in Finland (spruce and pine forests) where six plots of varying density were scanned at each. A program was written to automatically remove falling snow and noise from the point clouds and then estimate hemispherical gap fraction from any point. Traditional hemispherical photographs were taken in the same plots.

As a first test of the lidar's ability parameterise an RT model, gap fraction estimates from lidar and hemispherical camera were compared. The sensitivity of the lidar estimates to the cleaning algorithm, number of scans used and weather conditions (wind and snowfall) and of the hemispherical camera estimates to thresholding, gap joining (a post processing step) and illumination conditions were explored. The results suggest that terrestrial laser scanning is a powerful tool for parameterising RT models, however, care must be taken to avoid the blurring effect of the laser beam width and to cope with gaps between returns on solid objects.

This is part of a NERC funded project, grant number NE/H008187/1, which aims to improve the representation of how trees affect the shortwave and longwave radiation reaching the snow covered ground in land surface models. This should help improve the accuracy of weather and climate predictions.

References

Rutter et al, 2009. Evaluation of forest snow processes models (SnowMIP2). Journal of Geophysical Research, 114, D06111.

Parameterizing Climate Sensitive Forest Growth Models via LiDAR and Satellite Remote Sensing Data

Michael J. Falkowski[1], Andrew T. Hudak[2] & Linda M. Nagel[3]

[1]School of Forest Resources and Environmental Science, Michigan Technological University, mjfalkow@mtu.edu
[2]USDA Forest Service, Rocky Mountain Research Station, Forestry Sciences Laboratory, ahudak@fs.fed.us
[3]School of Forest Resources and Environmental Science, Michigan Technological University, lmnagel@mtu.edu

Paper Number: SL2012-131

Abstract

A detailed understanding of how forest composition, structure, and function will be impacted by projected climate change and related adaptive forest management activities are particularly lacking at local scales, where on-the-ground management activities are implemented. Climate sensitive forest growth models may prove to be effective tools for developing a comprehensive understanding. However, to be applicable to both regional forest planning and operational forest management, modeling approaches must be capable of simulating forest dynamics across large spatial extents (required for regional planning) while maintaining a high-level of spatial detail (required for operational management). LiDAR remote sensing has shown great utility for operational forest management, including forest growth modeling, albeit across relatively small spatial extents. We present a geospatial modeling approach to spatially initialize two separate climate-sensitive forest growth models (Climate-FVS and LANDIS-II) across unique ecoregions (in terms of forest structure and composition) in the Pacific Northwest of the US via an integration of sub-orbital LiDAR data with satellite remote sensing data. The system provides detailed forest inventory information - at both the landscape and ecoregion level - that is subsequently employed to demonstrate how climate sensitive growth and yield models can be used to 1) investigate the potential impacts of climate change on future forest composition and structure, and 2) assess how various forest management practices may either enhance or degrade forest resilience to changing climate and disturbance regimes.

Stand Structure Classificiation Using Airborne Laser Scanning (ALS) Data

Ian Moss[1] & Valerie LeMay[2]

[1]Tesera Systems, Victoria, BC
ian.moss@tesera.com
[2]Faculty of Forestry, University of British Columbia, Vancouver, BC
Valerie.LeMay@ubc.ca

Paper Number: SL2012-137

Abstract

Stand structure is the vertical and horizontal variation of trees in stands. In this study systems of stand structure (ground plot) classification are constructed using fuzzy c-means clustering according to similarities and differences in cumulative distributions of both stems per hectare and basal area per hectare, with respect to diameter at breast height.

The focus of this paper is on the reliability of ALS data for classifying stand structure with and without ground plot calibration. The study area was located in the southwest of Alberta near the eastern boundary of Banff National Park. The study area covered 26,444 ha of forested land. The forests are dominated by lodgepole pine (Pinus contorta var. latifolia Dougl. ex. Loud., Pl), white spruce (Picea glauca (Moench) Voss, Sw), and aspen (Populus tremuloides Michx., Aw). These and other species often grow as mixed-species stands, resulting in both vertical and horizontal complexity.

This investigation addressed the following questions: (1) How well-correlated are ALS and ground-level data? (2) Can pre-existing ground classes be distinguished according to selected ALS variables? (3) If new systems of classification are developed using ALS data alone, how well do these ALS classes relate to distributions of both stems and basal area per ha by DBH for determining ground-level stand structure class similarities and differences? How does this compare with the use of distributions of other ground variables, such as volume with respect to DBH or alternatively height? (4) If new systems of classification are developed based on ground measured distributions, how well do these ground classes relate to similarities and differences in distributions of ALS data, with and without the use of Multiple Discriminant Analysis (MDA)? (5) How does the number of stand structure classes affect classification results (re: questions 3 and 4)?

ALS data were collected using an Optech 3100 system on September 13, 2005 in the northern portion of the study area, and between June 29 and July 4, 2006 in the southern portion of the area. One hundred and eighty nine, 10 m radius ground plots were established to represent the range of stand conditions. Cumulative distributions in the proportions of first and last returns with respect to height were used to represent the ALS data. ALS performed reasonably well as a classifier of structure with (KHAT=0.591; 5 classes) or without (0.536) a priori ground plot calibration when evaluated according to cumulative distributions in volumes versus dbh. However it performed less well when classification was evaluated in terms of cumulative distributions in stems per hectare, with (0.421) or without (0.346) ground plot calibration. Classification success follows a negative exponential pattern with an increasing number or classes.

Biomass change estimation within NASA's Carbon Monitoring System

Jacqueline Rosette[1,2], Bruce Cook[1], Ross Nelson[1], Chenquan Huang[2], Jeff Masek[1], Compton Tucker[1], Guoqing Sun[1,2], Wenli Huang[2], Paul Montesano[1], Jérémy Rubio-Gil[1,2] and Jon Ranson[1]

[1]NASA Goddard Space Flight Center, Biospheric Sciences Laboratory, Greenbelt, MD 20771, USA jacqueline.rosette@nasa.gov, bruce.cook@nasa.gov, ross.f.nelson@nasa.gov, jeffrey.g.masek@nasa.gov, compton.j.tucker@nasa.gov, guoqing.sun-1@nasa.gov, paul.m.montesano@nasa.gov, jeremy.rubio-1@nasa.gov, kenneth.j.ranson@nasa.gov

[2]Department of Geographical Sciences, University of Maryland, College Park, MD, 20742, USA cqhuang@umd.edu, wlhuang@umd.edu

Paper Number: SL2012-141

Abstract

This research forms part of NASA's Carbon Monitoring System (CMS) initiative. One component of CMS is the local-scale mapping of biomass and stock change, including error budgets, at a scale that is ecologically meaningful and relevant for forest management. Given the availability of appropriate datasets, the aim is to explore methodologies for estimating biomass change using field survey, lidar analysis and optical remote sensing techniques.

The Howland Experimental Forest, Maine, USA was established as a research site in 1986 and covers an area of approximately 7000 hectares. The forest is retained without management intervention and is considered representative of boreal transitional forest. It is located at 45.2º N, 68.73º W and has relatively flat terrain at an elevation of 60 metres. Forest composition consists of approximately 91% conifers and 9% broadleaf species of varying ages.

Throughout the forest, several remote sensing datasets and products were utilised. Medium footprint, full waveform lidar data from NASA's Laser vegetation Imaging Sensor (LVIS) are available for 2003 and 2009. Small footprint discrete return lidar data were collected during September 2009 and the forest is due to be re-flown during Summer 2012 with NASA Goddard's Lidar Hyperspectral and Thermal instrument (GLiHT).

Using Jenkins generalised US biomass equations, biomass was estimated for 164 sub-plots within 12 1ha field plots, measured in 2009-2010 and dispersed throughout the forest. Relationships between field survey calculations of biomass and 2009 lidar metrics enabled biomass to be estimated throughout Howland Forest. Repeat lidar datasets from LVIS in 2003 and the planned G-LiHT campaign in 2012 allow(ed) forest height growth to be observed. These height increments are related to biomass change, thereby permitting a time series of biomass change to be produced through periodic lidar observations.

Independent datasets and methodologies will allow an assessment to be made of estimated biomass change. These sources include:

1. Field measurements of a 200m by 150m research area of Howland Forest which was established in 1989 and updated through complete re-measurement (2003) and partial re-measurement (2006 and 2010)
2. The Landsat-derived Vegetation Change Tracker disturbance product which estimates vegetation age, approximated as time since disturbance. Relationships between age and biomass allow growth curves to be established for the forest, substituting space for time, and for regions of disturbance between lidar acquisitions to be identified.
3. Carbon sequestration estimates from flux tower measurements.

Additionally, biomass estimation and change detection capabilities of a fine resolution and high

density airborne lidar sensor are compared with those of the coarser resolution, medium footprint spaceborne-lidar simulator, LVIS.

Virtual Forests for Developing Retrieval Algorithms of LiDAR Canopy

Sylvain G. Leblanc[1], Richard A. Fournier[2]

[1] Centre d'Applications et de Recherches en Télédétection, Department of Applied Geomatics, Université de Sherbrooke, Sherbrooke (QC) Canada J1K 2R1
SGLeblanc@videotron.ca

[2] Centre d'Applications et de Recherches en Télédétection, Department of Applied Geomatics, Université de Sherbrooke, Sherbrooke (QC) Canada J1K 2R1
Richard.Fournier@USherbrooke.ca

Paper Number: SL2012-143

Abstract

We propose a new method to estimate Leaf Area Index (LAI) and Plant Area Index (PAI) from LiDAR data using, as a baseline, methods developed for digital hemispherical photography (DHP). We combined a canopy architecture model with three remote sensing techniques: DHP and terrestrial and airborne laser scanning (ALS and TLS). The architectural model created 3D virtual forest environments based on statistical distributions of canopy components. Trees stems were distributed in rows (plantations), random, or clumped in groups. The foliage was simulated with random, erectophile or planophile leaf distributions, and the leaves were either clumped on branch structures or randomly distributed in crown shapes, such as spheres, cones, or cylinders. The 3D simulations were made using the freeware ray tracer POV-RAY that produces photorealistic renditions of complex scenes. The LiDAR data was simulated with the ground, leaves, and woody material in the scene colour-coded with height (z) and the known view geometry was used to get the horizontal coordinates (x-y). Simulations were made using the beam density, sampling configuration, and post-processing of different TLS and ALS sensors.

We used the LAI retrieval from DHP methods as our baseline. Since the foliage is not always separable from the wood, the PAI was retrieved first. A new PAI/LAI algorithm was developed for TLS data using a technique similar to the DHP method CLX, but for which the canopy element height information is used. The new algorithm was tested using both TLS and ALS. Seventy-five virtual forest sites were created, each with ten DHP, ten TLS scans, and ALS simulations for twelve view angles covering an area of 30 × 40 m^2 each. DHP and TLS scans were simulated at exactly the same position and with compatible resolutions. The new TLS algorithm was used to estimate PAI for ten height levels using the foliage clumping index at each level. For the DHP simulations, we observed that estimating PAI/LAI using the CLX clumping algorithm gave the most consistent results with R^2 of 0.7 and RMSE of 1.0 LAI units. The new LiDAR algorithm gives an R^2 of 0.8 and RMSE of 0.7 LAI units. Using the TLS method with airborne simulations at a view zenith angle of 20 degrees gave better results by about 25% compared with those at near nadir. Furthermore, LAI/PAI estimates at 50 degrees gave better results by almost 50% compared with those at near nadir. The simulation results also showed that resolution and spacing needs to be considered when adapting the hemi-LiDAR algorithm to sparser and coarser airborne LiDAR data. There is a high correlation between ALS PAI at large view angles and DHP PAI with a R^2 of 0.9, indicating that all techniques may suffer from the same errors that can largely be attributed to very dense tree crowns.

Statistical properties of biomass estimators in multi-level sampling designs

Hans-Erik Andersen[1] & Steen Magnussen[2]

[1]USDA Forest Service Pacific Northwest Research Station, Seattle, WA, USA
handersen@fs.fed.us
[2]Canadian Forest Service, Victoria, BC, Canada
Steen.Magnussen@NRCan-RNCan.gc.ca

Paper Number: #SL-2012-153

Abstract

More precise estimates of biomass and aboveground carbon are required in the boreal forests of interior Alaska, to support more accurate carbon accounting and enable the development of sustainable bioenergy systems in rural communities. Due to the high cost (~$8000 USD/plot) and logistical complexity of establishing field plots in remote areas of interior Alaska, which is largely devoid of the transportation infrastructure that exists elsewhere in the United States, there is a need for sampling approaches that draw heavily on remote sensing data – collected at varying resolutions with varying costs – while minimizing the number of costly field plots. In this study, we develop the statistical properties (mean square error, or MSE) of model-based total biomass estimators in a multi-level sampling design which uses 1) a limited number (n=79) of field plots, 2) airborne lidar strip sampling, and 3) wall-to-wall multispectral and L-band PolSAR satellite imagery over a 200,000 hectare area in the upper Tanana valley of interior Alaska. This design incorporates elements of both multi-phase and multi-stage sampling, where regression is used to establish the relationship between field plots and lidar metrics, resulting in biomass predictions along the lidar strips, and k-NN imputation is used to link the biomass predictions within the lidar strips to multispectral reflectance and L-band radar backscatter within imagery covering the entire study area. In addition, small area estimates of biomass resources available to support bioenergy production are obtained using higher-resolution (Quickbird) satellite imagery over a limited area surrounding the rural communities of Tok and Tanacross. Given the relatively high degree of complexity that characterizes the error structure of these multi-level designs, as well as the impressive capabilities of modern computing systems, resampling (i.e. bootstrap) and Monte Carlo simulation approaches to estimation of MSE provide a preferable alternative to analytical formulations. In this study, a double-bootstrap approach is used to estimate the model-based MSE for this multi-level design that accounts for variability introduced at every level in the design, including field sampling error, lidar strip sampling error, and modeling error in the imputation step (via parametric bootstrap).

References

Andersen, H.-E., J. Strunk, H. Temesgen, D. Atwood, and K. Winterberger. 2011. Using multilevel remote sensing and ground data to estimate forest biomass resources in remote regions: a case study in the boreal forests of interior Alaska. *Canadian J. of Remote Sensing* 37(6): 596-611.

Ståhl, G., S. Holm, T. Gregoire, T. Gobakken, E. Næsset, and R. Nelson. 2011. Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway. *Canadian J. of Forest Research* 41(1):96-107.

DART-LiDAR Model Simulations of Full-Waveform and Photon Counting Instruments

Jérémy Rubio[1,2,3], Eloi Grau[2], Jean Philippe Gastellu Etchegorry[2], Bruce Cook[3], Doug Morton[3] and Guoqing Sun[1,3]

[1] ESSIC, University of Maryland, College Park, MD, 20742, USA
rubio.jeremy@gmail.com
[2]CESBIO, 18 ave Edouard Belin, bpi 2801, 31401 Toulouse Cedex 9, France
eloi.grau@cesbio.cnes.fr, jean-philippe.gastellu-etchegorry@cesbio.cnes.fr
[3]NASA GSFC, Biospheric Sciences Laboratory, Greenbelt, MD 20771, USA
bruce.cook@nasa.gov, douglas.morton@nasa.gov, guoqing.sun-1@nasa.gov

Paper Number: SL2012-154

Abstract

NASA has developed and deployed a broad range of airborne and space-based LiDAR instruments for remote sensing of the Earth's surface. Instruments on ICESat and ICESat-2 use two fundamentally different LiDAR technologies, full-waveform and photon-counting, yet the comparability of measurements from these instruments has not been assessed. The goal of this study is to develop a theoretical basis for evaluating and comparing measurements of Earth surface properties from full-waveform and photon-counting LiDAR technologies. A LiDAR module was recently developed for the Discrete Anisotropic Radiative Transfer (DART) model. DART provides an ideal foundation for LiDAR modelling, because the architecture supports Monte Carlo and ray tracing methods to simulate radiative transfer in complex 3D scenes, thereby enabling realistic simulations of spaceborne systems.

We present results from DART-LiDAR model simulations for both full waveform and photon-counting instruments. Field measurements from a stem mapped study site in Howland Forest, Maine, USA, were used as input to the model. Model performance was verified and validated with independent airborne observations from NASA's Laser Vegetation Imaging Sensor (LVIS). A sampling approach for simulating data was also developed based on instrument specifications for the ICESat-2 mission. A sensitivity analysis was performed with the model to evaluate instrument characteristics (e.g., laser wavelength, detector response), canopy and ground surface properties (e.g., LAI leaf angle distribution, reflectance), and site conditions (e.g., presence of snow, wet/dry litter, and terrain slope) on the ability to characterize vegetation structure.

In the future, an atmospheric component of DART will be incorporated into the model and validated at sites where ICESat-2 airborne simulator data have been acquired. We will also evaluate the impacts of atmospheric contamination (e.g., biomass burning aerosols) on space-based observations of forest ecosystems. These results and the future development of DART-LiDAR model represent a significant advance in the ability to simulate the performance of spaceborne LIDAR systems, since there is currently no radiative transfer model that is capable of simulating both full-waveform and photon counting LiDAR returns that couples atmospheric, vegetation and surface interactions.

Simulation of micro-pulse photon counting lidar using the FLIGHT radiative transfer model for SIMPL and ICESat-2

Peter North[1], Jacqueline Rosette,[1,2,3], David Harding[2], Bruce Cook[2], Philip Dabney[2], Susan Valett[2], Jess Parker[4] and Josh Brinks[4]
[1] Swansea University, Department of Geography, UK, p.r.j.north@swan.ac.uk
[2] NASA Goddard Space Flight Center, Biospheric Sciences Laboratory, Code 618, Greenbelt, Maryland 20771, USA.
[3] University of Maryland College Park, Department of Geographic Sciences, College Park, Maryland, 20742, USA.
[4] Smithsonian Environmental Research Center, Edgewater, Maryland 21037, USA.

Paper Number: SL2012-160

Abstract

We present results from development and testing of a radiative transfer model to simulate new micro-pulse photon counting lidar sensors over vegetated terrain. The aim is to relate returned signal to biophysical parameters such as forest height and sub-canopy topography, and to prepare for future missions such as ICESat-2, scheduled for launch in 2016. The model is developed from the FLIGHT radiative transfer code, based on Monte Carlo simulation of photon transport (North 1996, North et al 2010). The model represents multiple scattering of light, and predicts the probability density function of photon return at a range of heights. LiDAR pulse, topography, solar radiation, atmospheric attenuation and spatial and temporal sampling of the instrument are explicitly modelled. The model is tested by comparison with observations from the Slope Imaging Multi-polarization Photon-counting Lidar (SIMPL). This is an airborne laser altimeter developed through the NASA Earth Science Technology Office Instrument Incubator Program (Harding et al, 2011). SIMPL incorporates beam splitting, single-photon ranging and polarimetry technologies at green and near-infrared wavelengths in order to achieve simultaneous sampling of surface elevation, slope, roughness and scattering properties. SIMPL data were captured for the Smithsonian Environmental Research Center (SERC), Edgewater, Maryland, during July 2010. Within a 400 x 400 m area of this forest, the locations and diameter at breast height (DBH) of approximately 33,400 trees (with DBH>1cm) were measured between February 2008 and November 2011. Locally-validated allometric equations were applied to reconstruct tree height and crown dimensions for the area. Eight SIMPL flightlines were selected to represent areas with different species composition, tree density and habitat conditions. The aim is to validate model simulations and to investigate from the effect of instrument configuration on retrieval of vegetation and surface properties.

References

Harding D, Dabney P, Valett S., 2011. Polarimetric, two-color, photon-counting laser altimeter measurements of forest canopy structure. In: International Symposium on Lidar and Radar Mapping - Technologies and Applications; May 26-29, 2011; Nanjing, Peoples Republic of China. 2011. 828629

North, P.R.J., 1996. Three-Dimensional Forest Light Interaction Model Using a Monte Carlo Method. IEEE Transactions on Geoscience and Remote Sensing, 34(4): 946-956.

North, P.R.J., Rosette, J.A.B., Suárez, J.C. and Los, S.O., 2010. A Monte Carlo radiative transfer model of satellite waveform lidar. International Journal of Remote Sensing 31(5): 1343-1358.

Combined Three-Dimensional Canopy Structure and Spectral Data for Fuels Mapping in Sagebrush-Steppe

Glenn, Nancy F.[1], Mitchell, Jessica[2], Shrestha, Rupesh[3] & Spaete, Lucas[4]

[1]Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory, Boise, ID, USA glennanc@isu.edu
[2]Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory, Boise, ID, USA, mitcjess@isu.edu
[3]Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory, Boise, ID, USA, shrerupe@isu.edu
[4]Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory, Boise, ID, USA, spaeluca@isu.edu

Paper Number: SL2012-167

Abstract

In the past 15 years, the western U.S. has experienced record fire activity, in terms of size, frequency, and intensity. Fire in the Great Basin sagebrush-steppe ecosystem has resulted in cheatgrass invasion across millions of acres, fragmentation of wildlife habitat, and erosion. Accurate spatially explicit fuel maps are essential inputs for fire prediction modeling and post-fire treatment across spatial and temporal scales. Fusing spectral time series data with three-dimensional vegetation structure data derived from laser altimetry (airborne or ground-based) increases the feasibility of mapping biomass dynamics in sparse, low-height vegetation ecosystems, such as the sagebrush-steppe. In this study, we develop methods to incorporate spectral data with 3-dimensional vegetation structural data to improve vegetation cover mapping and estimate biomass across the landscape. The vegetation cover mapping provides fuel stratum information, while the structural data provides density, and biomass (fuel) estimates needed to estimate fuel loads. Together, this information captures both variability in fuel structure and continuity (temporal and spatial) for improved fuels mapping.

Individual snag detection using airborne lidar data and 3D local-area point-based intensity filtration.

Brian M. Wing[1], Martin W. Ritchie[2], Kevin Boston[3], Warren B. Cohen[4], Michael J. Olsen[5]

[1] Oregon State University – College of Forestry brian.wing@oregonstate.edu
[2] US Forest Service – PSW Research Station mritchie@fs.fed.us
[3] Oregon State University – College of Forestry kevin.boston@oregonstate.edu
[4] US Forest Service – PNW Research Station warren.cohen@oregonstate.edu
[5] Oregon State University – College of Engineering michael.olsen@oregonstate.edu

Paper Number: SL2012-171

Standing dead trees (snags) play many essential roles in forested ecosystems at various scales. This has resulted in forest regulatory bodies and forest certification programs creating snag stocking standards to help ensure biodiversity is maintained or restored over time. The quantification and monitoring of snags across the landscape has been difficult due to their irregular and sparse distribution, often requiring intensive sampling methods to obtain statistically significant estimates. This study presents a new method for quantifying and locating snags using airborne discrete-return lidar. The semi-automated method uses both 2D and 3D local-area lidar point filters focused on spatial location and intensity information to identify individual lidar points associated with snags and eliminate points associated with live trees. The end result is a stem map of individual snags across the landscape with height estimates for each snag. The method was tested in four different forest types with similar results. Detection rates ranged from 35-100%, increasing as the size of snags increased. The overall detection rate for snags DBH $\geq$ 38 cm was 70.6% ($\pm$ 2.7%), with low commission error rates in all forest types. This information could be used to analyze the spatial distribution of large snags over entire landscapes, provide a better understanding of wildlife snag use dynamics, assess achievement of stocking standard requirements, and create more accurate snag stocking standards.

Autonomously operating terrestrial laser scanner for monitoring forest ecosystems at a very high temporal resolution

Jan U.H. Eitel[1,2], Lee A. Vierling[1,2], Troy S. Magney[1]

[1] Geospatial Laboratory for Environmental Dynamics, University of Idaho, Moscow, ID 83844-1135, USA, e-mail: jeitel@uidaho.edu (Eitel), leev@uidaho.edu (Vierling), tmagney@uidaho.edu (Magney).
[2] McCall Outdoor Science School, University of Idaho, McCall, ID 83638, USA

Paper Number: SL2012-176

Abstract

Light detection and ranging is a well-established technique for quantifying the structural properties of vegetation canopies. A major limitation of terrestrial laser scanning is the low temporal resolution that often exists between each data acquisition. The dynamic nature of forest ecosystems often requires datasets with high temporal resolution (e.g., daily data acquisition). The objectives of this study were to i) assemble an autonomously operating terrestrial laser system (ATLS) and to ii) to test its ability to map plant structural properties. We derived the height of Engelmann spruce (*Picea engelmannii*) saplings, the height of a graminoid (*Calamagrostis x acutiflora*) and forb (*Hemerocallis lilioasphodelus*), and DBH of Ponderosa Pine (*Pinus ponderosa*) and Douglas-fir (*Pseudotsuga menziesii*) trees from ATLS data and compared them with manual measurements. Our results showed a strong correlation between manual and ATLS derived tree sapling height (r^2 = 1.0, RMSE = 0.96 cm, slope = 0.96, intercept = 1.43), graminoid and forb height (r^2 = 0.98, RMSE = 1.56 cm, slope = 1.04, intercept = -2.22), and DBH (r^2 = 0.99, RMSE = 2.24 cm, slope = 0.99, intercept = 0.45). Our results indicate the suitability of ATLS to monitor forest structure at high temporal resolution. This type of information has great promise to provide new insights into forest dynamics.

Contrasting patterns of damage and recovery in logged Amazon forests from small footprint LiDAR data

D.C. Morton[1], M. Keller[2,3], B.D. Cook[1], M. Hunter[4], M. Sales[5], L. Spinelli[3], D. Victoria[3], H.-E. Andersen[6], S. Saleska[7]

[1]NASA Goddard Space Flight Center, Greenbelt, MD, USA, douglas.morton@nasa.gov, bruce.cook@nasa.gov
[2]USDA Forest Service, International Institute of tropical Forestry, San Juan, Puerto Rico, USA, mkeller.co2@gmail.com
[3]EMBRAPA Satellite Monitoring, Campinas, São Paulo, Brazil, daniel@cnpm.embrapa.br, spinelli@cnpm.embrapa.br
[4]University of New Hampshire, Durham, NH, hunter.maria@gmail.com
[5]Imazon, Belém, Pará, Brazil, marcio@imazon.org.br
[6]USDA Forest Service, Pacific Northwest Research Station, Seattle, WA USA, handersen@fs.fed.us
[7]University of Arizona, Department of Ecology and Evolutionary Biology, Tucson, AZ, USA, saleska@email.arizona.edu

Paper Number: SL2012-179

Abstract

Tropical forests ecosystems respond dynamically to climate variability and disturbances on time scales of minutes to millennia. To date, our knowledge of disturbance and recovery processes in tropical forests is derived almost exclusively from networks of forest inventory plots. These plots typically sample small areas (≤1 ha) in conservation units that are protected from logging and fire. Amazon forests with frequent disturbances from human activity remain under-studied. Ongoing negotiations on REDD+ (Reducing Emissions from Deforestation and Forest Degradation plus enhancing forest carbon stocks) have placed additional emphasis on identifying degraded forests and quantifying changing carbon stocks in both degraded and intact tropical forests. We evaluated patterns of forest disturbance and recovery at four ~1000 ha sites in the Brazilian Amazon using small footprint LiDAR data and coincident field measurements. Large area coverage with airborne LiDAR data in 2011-2012 included logged and unmanaged areas in Cotriguaçu (Mato Grosso), Flona do Jamari (Rondônia), and Floresta Estadual do Antimary (Acre), and unmanaged forest within Reserva Ducke (Amazonas). Logging infrastructure (skid trails, log decks, and roads) was identified using LiDAR returns from understory vegetation and validated based on field data. At each logged site, canopy gaps from logging activity and LiDAR metrics of canopy heights were used to quantify differences in forest structure between logged and unlogged areas. Contrasting patterns of harvesting operations and canopy damages at the three logged sites reflect different levels of pre-harvest planning (i.e., informal logging compared to state or national logging concessions), harvest intensity, and site conditions. Finally, we used multi-temporal LiDAR data from two sites, Reserva Ducke (2009, 2012) and Antimary (2010, 2011), to evaluate gap phase dynamics in unmanaged forest areas. The rates and patterns of canopy gap formation at these sites illustrate potential issues for separating logging damages from natural forest disturbances over longer time scales. Multi-temporal airborne LiDAR data and coincident field measurements provide complementary perspectives on disturbance and recovery processes in intact and degraded Amazon forests. Compared to forest inventory plots, the large size of each individual site permitted analyses of landscape-scale processes that would require extremely high investments to study using traditional forest inventory methods.

Information Fusion for Automatic Delineation of Individual Tree Crowns from Multi-source Remote Sensing Data

Baoxin Hu[*1], J. Li[1], and L. Jing[2]

[1]Department of Earth and Space Science and Engineering, York University, Toronto, Ontario, Canada M3J 1P3. E-mails: baoxin@yorku.ca, ljili@yorku.ca
[2]Center for Earth Observation and Digital Earth, Chinese Academy of Sciences, China, 100094. Email: jinglinhai@hotmail.com

Paper Number: SL2012-182

Abstract

Individual tree crown (ITC) delineation is a very important step in facilitating aspects of sustainable forest management. ITCs serve as the basic unit for species identification, gap analysis, and volume or biomass estimation, etc. In recent decades, various methods of ITC delineation using either very high spatial resolution (passive) optical imagery or LiDAR (Light Detection And Ranging) data have been developed. Even though some promising results have been reported in the literature for simple structured un-natural or coniferous forests, the delineation accuracy remains low for natural deciduous or mixed forests. With optical imagery, tree crown delineation is negatively affected by the anisotropic reflectance properties of tree canopies, which is especially significant for the open canopies. In addition, understory vegetation can be falsely accounted as trees. On the other hand, a tree crown tends to have distinct boundaries in the 3-dimension (3D) data clouds collected by a LiDAR instrument. However, for close canopies, trees grow together, which makes the distinction between crowns in the LiDAR data less evident. In this study, the spectral information derived from optical imagery, and height and architectural (such as stem) information from LiDAR data clouds were employed intelligently to improve ITC delineation for natural forests.

With the developed method, individual tree tops were first located from the optical imagery and LiDAR data separately using a novel multi-scale CSI (Crown Slice from Imagery) method proposed in a previous study of ours (Jing et al., 2012). The detected tree tops were then integrated together. Using the detected tree tops, the marker-control watershed segmentation method was used to delineate individual tree crowns from optical imagery data. For each resulting segment, a tree stem(s) was detected from the LIDAR points within this segment. A segment could be kept as it was, split, or merged depending on the number of stems was detected.

In an experiment on natural forests in Ontario, Canada, the proposed method yielded crown maps having a good consistency with manual and visual interpretation.

Reference

Jing, L., Hu, B., Li, J., and Noland, T., 2012. Automated tree crown delineation from imagery based on morphological techniques. *IEEE Transactions on Geoscience and Remote Sensing*, submitted.

Simulating the impact of changing airborne lidar sensor and survey properties on estimation of vegetation structure parameters

John Armston[1,2], Mathias Disney[3], Peter Scarth[1,2], Richard Lucas[4], Stuart Phinn[1], Philip Lewis[3] & Nicholas Goodwin[2]

[1]Joint Remote Sensing Research Program, School of Geography, Planning and Environmental Management, University of Queensland
j.armston@uq.edu.au; p.scarth@uq.edu.au; s.phinn@uq.edu.au
[2]Remote Sensing Centre, Science Delivery, Department of Science, Information Technology, Innovation & the Arts, nicholas.goodwin@qld.gov.au
[3]Department of Geography, University College London
mathias.disney@ucl.ac.uk; p.lewis@ucl.ac.uk
[4]Institute of Geography and Earth Sciences Aberystwyth University, rml@aber.ac.uk

Paper Number: SL2012-188

Abstract

Rapid advances in airborne lidar systems have introduced inconsistency into vegetation structure parameter estimates derived from multi-temporal and large area datasets captured using different sensors. Recent comparisons of lidar products with ground measurements in Queensland, Australia, for calibration and validation of vegetation structure parameters have provided estimates of bias and variance across a range of environments however the relative importance of the sensor and survey properties on these error estimates was poorly understood. Details on commercial lidar systems are often undisclosed and coincident captures from different lidar sensors are often impractical. New approaches are required to understand the impacts of changing lidar sensor and survey properties on the estimation and monitoring of vegetation structure parameters.

We investigated the impact of airborne lidar sensor and survey properties on the estimation of gap probability profiles using Monte-Carlo ray tracing, geometrically explicit 3D canopy models, and Riegl LMS-Q680i waveform lidar data acquired at multiple flying heights over a eucalypt dominated savanna in north eastern Australia. The simulation approach was developed to quantify relative differences in the signal-to-noise of lidar waveforms caused by changes in laser pulse rate, range, and wavelength. The crown cover, gap fraction and foliage area volume density profiles of the 3D models showed agreement with independent field and ground-based lidar measurements at three test sites, with greatest differences in the lower canopy.

Comparison of measured and modelled gap probability profiles derived using waveform and discrete return methods and canopy/ground waveform backscatter coefficient ratios showed agreement, particularly in the shape of the profiles. Differences in magnitude were linked with bias in the within-crown gap probability of the 3D tree models. The modelled gap probability profiles replicated the direction and magnitude of bias observed in the Riegl LMS-Q680i estimated profiles for different flying heights and pulse rates. Simulation experiments at multiple pulse wavelengths showed that estimates of gap probability profiles using a new waveform model were less sensitive than discrete return methods to decreases in the signal-to-noise ratio caused by interrelated changes in footprint size, pulse energy, and range.

The results indicate the modelling approach presented in this study is able to replicate relative changes in gap probability profiles for a range of lidar sensor and survey properties in a complex eucalypt savanna. The approach is currently being used to investigate the impact of canopy and grass senescence on pulse-canopy interactions, and the detection of change in savanna gap fraction and derived vegetation structure parameters.

Mapping and Managing the Fuzzy Forest

Günther Bronner

Umweltdata GmbH Mödling, Austria; g.bronner@umweltdata.at

Paper Number: SL2012-193

Abstract:

Forest Management Plans (FMPs) are composed from tables, reports and maps. Mapping the forest usually means to segment the forest area into vector polygons. The tradition of drawing lines on forest land to separate different stands goes back to the roots of FMP in the 18th century. With high sophisticated GIS methods like ecognition® it has been brought to perfection within the last decade. But does it really make sense to split the forest into hundreds and thousands of vector polygons? Do we need this kind of generalization in our management plans? Forest inventory specialists would deny the need of segmented small units. With their statistical approach, they can perfectly answer the questions of stock volume, annual growth, annual cut rates and long-term sustainability-figures for larger units. Anyhow, the forest ranger wants to get rid of some manageable units. And we can take for sure: he will need maps for his day-to-day business. In many European regions, the commercial forestry has developed from a clear-cut-based tradition to different kinds of selection-felling-methods. The application of such nature-near techniques results in 'Fuzzy Forests' after some decades, forest areas, which will persistently resist any kind of vector segmentation. Fuzzy Forests are a challenge in FMPs as well as in forest cartography. After realizing, that splitting Fuzzy Forests into small segments means barking up the wrong tree, we came to the conclusion, that Lidar-based forest assessment will perfectly meet our requirements. After mapping 300.000 hectares of forest land with airborne LIDAR data, I will present different kinds of successfully proven thematic maps (fuzzy or segmented), discuss the methods used in production, highlight the usability and the feedback from forest rangers and give an outlook, how such geo-data will generate added-value as interactive FMPs on tablet PCs.

Forest Aboveground Biomass Estimation using ICESat GLAS and Imagery Remote Sensing Data in the Greater Mekong Subregion

Pang Yong[1], Li Zengyuan[1], Huang Kebiao[2]
[1]Institute of Forest Resource Information Technique, Chinese Academy of Forestry, Beijing 10009, China, caf.pang@gmail.com, zy@caf.ac.cn
[2]Asia-Pacific Network for Sustainable Forest Management and Rehabilitation, Beijing 100013, China, huang_kebiao@apfnet.cn

Paper Number: SL2012-195

Abstract

Forests play an irreplaceable role in maintaining regional ecological environment, global carbon balance and mitigating global climate change. Forest aboveground biomass (AGB) is an important indicator of forest carbon stocks. Estimating forest aboveground biomass accurately could significantly reduce the uncertainties in terrestrial ecosystem carbon cycle. The Greater Mekong Subregion (GMS) is rich in forest resources, the change of forest resources affect the regional even global climate change. It is important to estimate forest AGB with high accuracy methods in this region. Remote sensing is an efficient way to estimate forest parameters in large area, especially at regional scale where field data is scarce. LIDAR provides accurate information on the vertical structure of forests. Combining airborne LiDAR and spaceborne LiDAR for regional forest biomass retrieval could provide a more reliable and accurate quantitative information in regional forest biomass estimate. In this study, vertical structure of forest parameters of two forest farms in Yunnan province, China were derived using airborne LiDAR system (ALS). Regression models were built between field data of forest AGB and percentiles of canopy height, canopy density derived from ALS point cloud data. The high accuracy ALS estimated and field measured forest AGB were used as training data for building forest AGB estimation model with ICESat GLAS waveform indices. Then the forest ABG was estimated at ICESat GLAS footprint level in GMS. According to different types of ecological zones, a set of categorical regression models was built between ICESat GLAS estimates and MERIS spectral variables. Then, a forest aboveground biomass map with continuous biomass values was generated. The results showed that: 1) The integrated method based on field measurements, airborne and spaceborne LiDAR data can be used to estimate forest aboveground biomass effectively. Ten estimation equations were built using regression decision tree method. The overall average error of the estimation models was 34 ton/ha, with a correlation coefficient of 0.7. 2) The estimation agreed well with FAO FRA 2010 report and other published results, and the average difference was 13.3%. 3) The total forest aboveground biomass in the Greater Mekong Subregion estimated based on remote sensing data was 6.27 billion tons which consisted of 71% evergreen broadleaf forest, 10% deciduous broadleaf forest, 16% evergreen coniferous forest and 3% mixed forest. 4) According to the total aboveground biomass map, Myanmar has the largest AGB in the region which accounted for 22% of the total regional biomass.

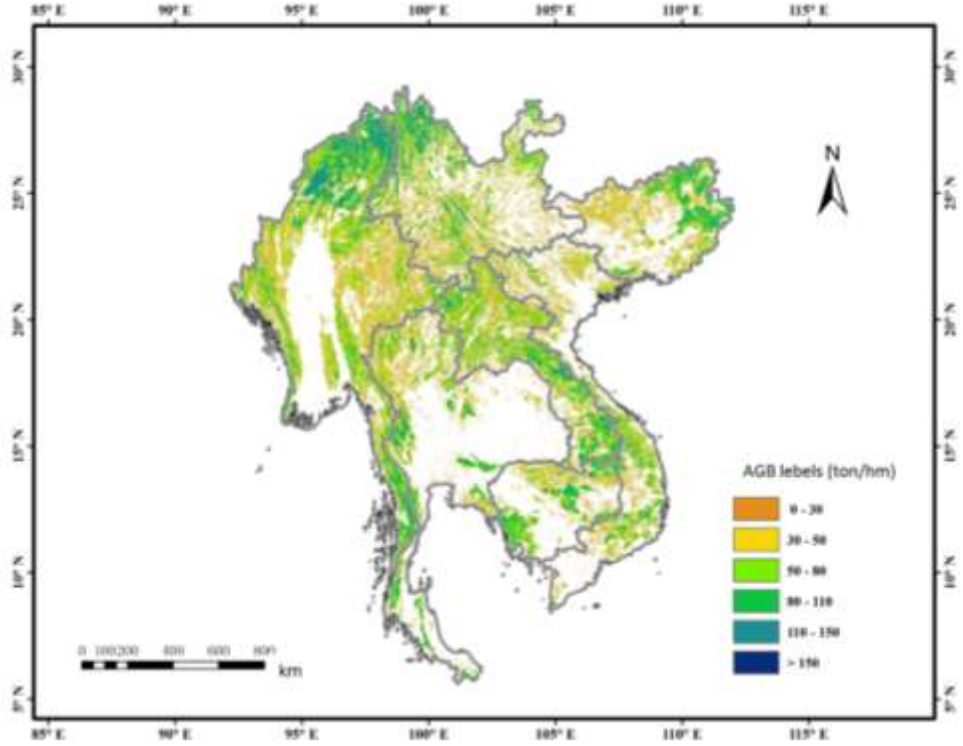


Figure 1. Estimated forest AGB in the Greater Mekong Subregion

Forest Structure and Biomass Mapping Using Time Series Landsat Observations, Small Footprint Lidar, and Field Inventory Data

Chengquan Huang[1], Yong Pang[1, 2], Chris Toney[3], Jeffrey Masek[4], Ralph Dubayah[1]

[1] Department of Geographical Sciences, University of Maryland, College Park, MD 20742, USA. cqhuang@umd.edu
[2] Forest Remote Sensing Lab, Chinese Academy of Forestry, Beijing 100091, China. caf.pang@gmail.com
[3] USDA Forest Service, Rocky Mountain Research Station, Missoula, MT 59808, USA. christoney@fs.fed.us
[4] Biospheric Sciences Lab, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. jeffrey.g.masek@nasa.gov

Paper Number: SL2012-202

Abstract

Although significant efforts have been devoted to deriving forest structure and biomass using optical remote sensing data, relationships between spectral data and forest structure were often very weak, because the former is typically not very sensitive to changes in forest structure. However, optical remote sensing images, especially those produced through a series of 6 Landsat instruments, have been available since 1972. This imagery record makes it possible to detect forest disturbance and calculate age since disturbance for up to four decades. Because forest age is often a good predictor of forest growth and yield, the age since disturbance information derived using time series Landsat observations has been found highly valuable for predicting forest height growth. For forests where no disturbance was observed by the Landsat systems since the 1970s, a spectral record of their growth in multiple decades with the knowledge that they are likely older than 40 years will allow better estimation of their structure. In this study, we develop an approach for mapping forest structure and biomass using time series Landsat observations. In this approach, forests are first divided into "young" forests that generated following disturbances recorded by the Landsat and "old" forests that did not experience stand-replacement disturbance event since the first available Landsat observation. Height and biomass for these two types of forest are modelled separately using empirical methods trained using lidar or field plot data. This approach is used to develop forest structure and biomass products over North Carolina.

LAI estimation using airborne lidar and DBInSAR data

Alicia Peduzzi[1], Randolph H. Wynne[2],Valerie A. Thomas[3],
Ross F. Nelson[4], James J. Reis[5], and Mark Sanford[6]

[1]USDA Forest Service, apeduzzi@fs.fed.us
[2]Virginia Polytechnic Institute and State University, wynne@vt.edu
[3]Virginia Polytechnic Institute and State University, thomasv@vt.edu
[4]NASA Goddard Space Flight Center, Ross.F.Nelson@nasa.gov
[5]Fugro Earth Data, jreis@earthdata.com
[6]Fugro Earth Data, msanford@earthdata.com

Paper Number: SL2012-207

Abstract

The objective of this study (Peduzzi et al. 2012) was to determine whether leaf area index (LAI) in temperate mixed forests is best estimated using multiple-return airborne laser scanning (lidar) data or dual-band, single-pass interferometric synthetic aperture radar data (from GeoSAR) alone, or both in combination. In situ measurements of LAI were made using the LiCor LAI-2000 Plant Canopy Analyzer on 61 plots (21 hardwood, 36 pine, 4 mixed pine hardwood; stand age ranging from 12-164 years; mean height ranging from 0.4 to 41.2 m) in the Appomattox-Buckingham State Forest, Virginia, USA. Lidar distributional metrics were calculated for all returns and for ten one meter deep crown density slices (a new metric), five above and five below the mode of the vegetation returns for each plot. GeoSAR metrics were calculated from the X-band backscatter coefficients (four looks) as well as both X- and P-band interferometric heights and magnitudes for each plot. Lidar metrics alone explained 69% of the variability in LAI, while GeoSAR metrics alone explained 52%. However, combining the lidar and GeoSAR metrics increased the R^2 to 77% with a CV-RMSE of 0.42. This study indicates the clear potential for X-band backscatter and interferometric height (both now available from spaceborne sensors), when combined with small-footprint lidar data, to improve LAI estimation in temperate mixed forests. This result, evaluated in the context of demonstrated lidar-radar synergies for biomass estimation in both this study area (Banskota et al. 2011) and in ponderosa pine forests (Hyde et al. 2007), points to the need for further research into the combined use of lidar and radar data for forest biophysical parameter estimation.

References

Banskota, A., Wynne, R.H., Johnson, P., and Emessiene, B., 2011. Synergistic use of very high frequency radar and discrete-return lidar for estimating biomass in temperate hardwood and mixed forests. *Annals of Forest Science*, 68, 347-356.

Hyde, P., Nelson, R., Kimes, D., and Levine, E., 2007. Exploring lidar-radar synergy – predicting aboveground biomass in a southwestern ponderosa pine forest using lidar, SAR, and InSAR. *Remote Sensing of Environment*, 106, 28-38.

Peduzzi, A., Wynne, R.H., Thomas, V.A., Nelson, R.F., Reis, J.J., and Sanford, M., 2012. Combined use of airborne lidar and DBInSAR data to estimate LAI in temperate mixed forests. *Remote Sensing*, 4,1758-1780.

Operational implementation of airborne LiDAR for forest inventory attributes in Boreal Ontario – lessons learned and future directions.

Doug G. Pitt[1], Murray Woods[2] Margaret Penner[3], Kevin Lim[4], Paul Treitz[5]

[1]Canadian Wood Fibre Center, CFS dpitt@NRCan.gc.ca
[2]Ontario Ministry of Natural Resources murray.woods@ontario.ca
[3]Forest Analysis Ltd. mpenner@vianet.ca
[4]Lim Geomatics Inc. kevin@limgeomatics.com
Queen's University treitzp@queensu.ca

Paper Number: SL2012-209

Abstract

Two northeastern Ontario boreal forest areas, totaling 2 million hectares, were provided with *wall-to-wall* coverage of airborne LiDAR at approximately 0.5 to 1 pulse per m^2. Within each of the two areas, we defined standard forest types that were of interest to local managers and then sampled each of these, through the range of development stages present, with approximately thirty 400-m^2 field plots. Stand-level inventory attributes measured from these plots were regressed on selected LiDAR point-cloud statistics, by forest type, to produce parametric predictive models for basal area, average tree size (height, diameter, stem volume), stem density, volume (gross total, gross merchantable, and sawlog), above-ground biomass, and tree size distribution. In validation, these models revealed coefficients of variation for predicted versus observed values that were typically about 10% for average tree size variables, and about 20% for stand volume-type variables. The models were then used to produce 20-m × 20-m predictive rasters for the forest inventory variables across each of the forests. These rasters have proven highly useful for informing decisions and improving the efficiency of forest operations through better planning, optimum selection of harvesting equipment, and better allocation of raw materials to mill facilities. Assuming forest-type definition compatibility, predictive models may be transferable from one forest to the other, however, the need to validate such transfers and correct for bias may offset any potential cost savings. A disadvantage of the parametric model-building approach used is that it depends on the pre-existence of stratification by forest type (e.g., current forest resource inventory polygons), and that this stratification is accurate at the scale of the 20-m × 20-m LiDAR-based prediction unit. Increasingly popular non-parametric approaches (e.g., RandomForest) are allowing us to make predictions in the absence of such stratification. Alternatively, we have had early success in the prediction of forest type directly from LiDAR.

Challenges for Measuring and Monitoring Global Forest Carbon Stocks and Change Using Lidar Remote Sensing

Ralph Dubayah[1]& Scott Goetz[2]
[1]Geographical Sciences, University of Maryland, dubayah@umd.edu
[2]Woods Hole Research Center, goetz@whrc.org

Paper Number: ###SL2012-210

Abstract

Forest monitoring using airborne and satellite data has advanced tremendously over the past few decades, to the point that these datasets now inform international policy agreements, notably those associated with emissions of CO2 into the atmosphere from deforestation and other types of land-use change. However, technological advances require time to move towards a state of operational readiness for global monitoring and reporting and this has been the case with lidar remote sensing methods. While lidar has promised to revolutionize global forest carbon measurements, significant challenges persist. Among these are the large gap in capability between airborne and spaceborne lidar and the lack of stable methodological frameworks for estimating carbon stocks. The end result is the generation of products that often are inconsistent, inaccurate and whose errors are not well characterized. In this talk we explore the current and future status of lidar within the context of next-generation earth observation missions in support of science activities and international programs such as REDD. We will highlight areas where improvement is required and discuss potential approaches for regional and global monitoring. Such approaches most likely will be based on the deployment of new instruments, advances in multi-sensor fusion techniques and a much tighter coupling between structure observations and models. As part of a comprehensive suite of earth observation technologies, lidar is poised to play a definitive role but to do so will require attention to these issues and persistent advocacy for global measurement capabilities based on lidar and other active remote sensing.

References

Goetz, S and R Dubayah. Advances in remote sensing technology and implications for monitoring, reporting and verifying forest carbon stocks and emissions. *Carbon Management*, 2, 231-244 (2011).

MacArthur's Ladder: Advancing the role of LiDAR in assessing and conserving biodiversity

Lee A. Vierling[1], Kerri T. Vierling[2], Sebastián Martinuzzi[3], Jody Vogeler[4] & Andrew T. Hudak[5]

[1]Department of Forest, Rangeland, and Fire Sciences, University of Idaho, Moscow, ID 83844, USA leev@uidaho.edu
[2]Department of Fish and Wildlife Sciences, University of Idaho, Moscow, ID 83844, USA kerriv@uidaho.edu
[3]Department of Forest and Wildlife Ecology, University of Wisconsin, Madison, WI 53706, USA martinuzzi@wisc.edu
[4]Department of Forest Ecosystems & Society, Oregon State University, Corvallis, OR 97331, USA jody.vogeler@oregonstate.edu
[5]United States Forest Service Rocky Mountain Research Station, Moscow, ID 83843, USA ahudak@fs.fed.us

Paper Number: SL2012-211

Abstract

In the 1950s and '60s, Robert MacArthur and colleagues combined painstaking field data collection using ladders with elegant mathematical reasoning to establish fundamental ecological theory relating animal and plant species diversity with ecosystem vertical structure (*z*). Today's growing storehouse of LiDAR data provides significant opportunity to advance the next breakthroughs in understanding 4-dimensional (*x, y, z, t*) patterns in biodiversity, a high priority in our world of species decline and unprecedented environmental change. A new wave of study has begun to demonstrate that LiDAR-based structural variables provide unique information to quantify the presence/absence, species richness/diversity, and/or habitat quality of a growing number of taxonomic groups including plants, birds, mammals, fish, beetles, and spiders. While recent successes are highly promising, characterizing biodiversity is a challenging endeavour demanding the highest level of interdisciplinary communication and skill. In light of these challenges, we propose four recommendations for future effort to advance the role of LiDAR in assessing biodiversity: (1) *Integration and collaboration*, (2) *Focus on function*, and (3) *Technology transfer*, all of which should occur within (4) *Management-relevant contexts*. *Integration and collaboration* is needed to coordinate and synthesize LiDAR and other remotely sensed datasets with field-based species census data across multiple sites and scales, and to analyze these data in the context of established ecological theory. Doing so will move the science beyond proof-of-concept and aid in identifying emergent properties across diverse sets of ecosystems and conditions. A *focus on function* will help to transcend the limitations inherent to studying taxonomic diversity, allowing LiDAR-quantified ecosystem structure to be examined from a functional diversity perspective. This approach is likely to be more scalable and holistic than taxonomic approaches, and may lead to mechanistic understanding of structure-function relationships. *Technology transfer* of open source analysis tools is increasingly important as LiDAR datasets become more widely available across the globe, particularly in areas of high species richness and endemism where local knowledge can be essential for interpreting results. Finally, future analyses should occur in *management-relevant contexts* so that the outcomes can be applied immediately to address priorities of biodiversity conservation. Conducting such collaborative LiDAR-based efforts with the resourcefulness and creativity of MacArthur is likely to be helpful in determining biodiversity status and trends in our increasingly human-dominated planet.

Poster Abstracts

The potential of terrestrial laser scanning for the estimation of biomass in coppice-with-standard systems

Dominik Seidel[1&2], Katja Albert[1], Lutz Fehrmann[2] & Christian Ammer[1]
[1]Silviculture and Forest Ecology of the Temperate Zones, Faculty of Forest Science, University of Göttingen
[2]Chair of Forest Inventory and Remote Sensing, Faculty of Forest Science, University of Göttingen
dseidel@gwdg.de; kalbert@gwdg.de; lfehrma@gwdg.de;
christian.ammer@forst.uni-goettingen.de

Paper Number: 005

Abstract

Methods for estimating the biomass potential of dense coppice or coppice-with-standard forests in a fast and objective way are currently rare. We adapted existing methodical approaches for biomass estimations from terrestrial laser scanning developed for mature stands in order to perform single scan measurements of diameter at breast height in extremely dense coppice (~30,000 stems ha^{-1}). Tree stems and according diameters at breast height were automatically detected by an algorithm and were used as input for allometric regression models in order to estimate the dry weights of each tree. As a tribute to the dense stocking on the investigated stands our study plots were smaller (circular 12.56 m²) than in previous studies focusing on mature forests (e.g. Maas et al. 2008).
Results were found to be consistent with a mean absolute error (MAE) of 9.9 kg for the dry weight estimation per plot, which is equal to a relative error of 15%. With respect to the strongly reduced amount of field work the method is therefore of high efficiency. With the new approach reliable assessments of the bioenergy potentials become possible for coppice stands if based on sound sampling designs, e.g. regular sampling grids. Such fully automatic and non-destructive biomass estimations will play an important role in future resource management in order to support mitigating climate change.

References

Maas, H.G., Bienert, A., Scheller, S. and Keane, E. (2008): Automatic forest inventory parameter determination from terrestrial laser scanner data. *Int J Remote Sens* 29: 5, 1579-1593.

Influence of LiDAR data projection in DTM generation

Eric Bastos Görgens[1], Luiz Carlos Estraviz Rodriguez[1], Nicholas Coops[2], André Gracioso Peres da Silva[1], Carlos Alberto Silva[1]

[1]University of São Paulo, ESALQ, CMQ. Piracicaba, São Paulo, Brazil.
gorgens@usp.br, lcer@usp.br, andregracioso@gmail.com,
carlos_engflorestal@yahoo.com.br
[2]Department of Forest Resources Management, Faculty of Forestry, University of British Columbia, Vancouver, Canada. nicholas.coops@ubc.ca

Paper Number: #SL2012-018

Abstract

GPS measures elevations based on a mathematical representation of the Earth. The ellipsoid is usually used to represent the shape of the Earth, and its geocentric point serves as the reference to determine the elevation of each LiDAR return. WGS84 is the most common coordinate frame used by GPS equipments. In some occasions, LiDAR files may come with negative values for elevation, caused by differences between the real surface and the reference ellipsoid. This paper investigates how LiDAR data negative values affect the generation of DTM when the Kraus & Pfeifer algorithm is used and how to avoid mistakes in such cases. The creation of DTM from LiDAR dataset is based on the recognition of return points associated with ground level. This means to scan the LiDAR cloud and identify which return came from the interaction of the laser and the ground. The filter algorithm proposed by Kraus and Pfeifer in 1998 applies weights for each LiDAR point as a function of the distance for a surface reference. Based on those weights, a new surface is generated. Points below the reference surface should influence more than points above the reference surface. Interactively, a new surface is generated using the points which most probably represent the terrain. The study was based on data sampled from the coastal area of the Vancouver Island. To eliminate negative values an offset was applied to the data considering the minimum value. Both datasets, original and offset, were submitted to specific FUSION and Octave routines. The surface created in FUSION from the original data set missed all the negative values, and did not bound the ground points. In opposite, if the offset is applied to the data, eliminating the negative values, the surface bounds correctly. The Kraus and Pfeifer filter implemented in Octave showed that the algorithm works properly for both situations: negative and positive values. The result shows the distortion did not come from the filter procedure. Negative values arise because of the assumption that the shape of the Earth is a smooth ellipsoid. Consequently, natural variations on the real surface of the Earth can often generate negative elevations. To eliminate the effect of these negative values in the process of digitally modeling the terrain, an effective and simple approach is to apply the offsetting method based on the most negative observed value in the LiDAR dataset. Alternatively, another possible approach is to change the reference system used to represent the shape of the Earth.

HTML5 based Visualization of high density LiDAR data and color information for agriculture applications

Bo Mao[1,2], Jie Cao[1], Zhiang Wu[1] & Wei Yao[3,4]

[1] Jiangsu Provincial Key Laboratory of E-Business, Nanjing University of Finance and Economics, Nanjing, P.R. China, maoboo@gmail.com
[2] Geoinformatics, Royal Institute of Technology, Stockholm, Sweden
[3] Photogrammetry and Remote Sensing, Technische Universitaet Muenchen, 80333 Munich, Germany, yao@tum.de
[4] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications of Chinese Academy of Sciences, Beijing, China

Paper Number: SL2012-022

Abstract

In this paper, an online visualization framework for agriculture LiDAR point cloud is proposed based on HTML5 technologies. In the framework, 3D point datasets are transmitted with Websocket, and visualized with X3DOM that allows to integrate X3D elements as part of HTML5 DOM tree. No specific plugins for the browser are required by the framework, which extends the availability of the LiDAR data for agriculture related applications. The proposed online agriculture LiDAR point cloud visualization framework is illustrated in Fig. 1. First, the LiDAR dataset is transmitted into user browser by Websocket. Then the received point data is converted into DOM object in X3D format and directly updated in the HTML content. Finally, X3DOM will automatically visualize the 3D scenes in the created DOM objects based on WebGL. In the user client, if WebGL supported browser (such as Chrome, Firefox, Safari and Firefox mobile for android) is installed, the 3D point clouds will be visualize d by the proposed framework without any plugin. Next, we will introduce implementation details of the proposed framework.

Two types of LiDAR data on agriculture are studied in this paper. The proposed framework is implemented in LAN network environment. The server is Tomcat 7 running on a PC with Intel Core2 Quad 2.33GHz CPU, 3.00GB RAM and Microsoft Window XP SP3. The browser is Mozilla Nightly 13.0 running on a laptop with Intel Core2 Duo 2.40GHz CPU, 2GB RAM and Microsoft Window XP SP3. The development platform is Eclipse 3.7.1 running on the server. The experimental results demonstrate the effectiveness of the framework

Acknowledgements

This work is supported by Jiangsu Provincial Key Laboratory of E-Business, Nanjing University of Finance and Economics, Nanjing, P.R. China, NFS of China (71072172, 61103229), Jiangsu Province Colleges and Universities Outstanding Scientific and Technological Innovation Team Fund (No.2011013), Jiangsu Province Colleges and Universities Industrization of research findings promotion project (No. JHB2011-21), Transformation Fund for Agricultural S&T Achievements under Grants (No. 2011GB2C100024) and Innovation Fund for Agricultural S&T in Jiangsu under Grants No. CX(11)3039, Open Fund of State Key Laboratory of Remote Sensing Science (Grant No. OFSLRSS201212).

Detecting tree competition- from model to measurement

Dominik Seidel[1], Jérôme Metz[1] & Christian Ammer[1]
[1]Silviculture and Forest Ecology of the Temperate Zones, Faculty of Forest Science, University of Göttingen
dseidel@gwdg.de; jerome.metz@forst.uni-goettingen.de; christian.ammer@forst.uni-goettingen.de

Paper Number: 023

Abstract

The applicability of traditional yield tables and other simple, stand-orientated models to the forest of the future is limited by diversity in spatial structure, age and species composition, and especially by intra- and inter-specific competition and the resulting variety of growth patterns. The approximation of actual or available growing space is one important method for measuring competition in stands which led to the development of neighborhood models of plant performance. We evaluated the quality of the upside-down search-cone method for competitor election and competition quantification (e.g Rötzer et al. 2009) based on terrestrial laser scanning (TLS). The high resolution spatial information on tree structure obtained from multiple perspective TLS allowed us to evaluate whether traditional models of competitor election based on simple tree models for stand simulators (e.g. SILVA) are able to produce realistic estimations or not. An upside-down search cone was therefore inserted into a point cloud created from multiple-scan laser scanning in order to detect the competing trees of a target tree and to measure their competitive strength based on the amount of tree material reaching into the cone. For the first time existing theories of the detection of competition in forests were thereby evaluated based on real tree structure instead of simple tree models, with surprising results.

References

Rötzer, T., Seifert, T. and Pretzsch, H. (2009): Modelling above and below ground dynamics in a mixed beech and spruce stand influenced by climate. *European Journal of Forest Research* 128: 172-182.

Predicting Slope from Full Waveform LiDAR

Craig Mahoney[1], Natascha Kljun[1], Sietse Los[1], Peter North[1], Laura Chasmer[2], Eva van Gorsel[3], Jorg M. Hacker[4], Chris Hopkinson[2,5], and Jacqueline Rosette[1,5,6]

[1]Department of Geography, College of Science, Swansea Univesity, Swansea, SA2 8PP, international@swansea.ac.uk
[2]Wilfrid Laurier University, Waterloo, Ontario, Canada, webmaster@wlu.ca
[3]CSIRO, Marine and Atmospheric Research, Canberra, Australia, Enquiries@csiro.au
[4]Airborne Research Australia, Flinders University, Adelaide, Australia, info@airborneresearch.com.au
[5]NASA Goddard Space Flight Center, Biospheric Sciences Laboratory, Greenbelt, Maryland, USA, anetra.m.tucker@nasa.gov
[6]University of Maryland College Park, Department of Geographical Sciences, Maryland, USA, emailum@umd.edu

Paper Number: #SL2012-024

Abstract

We present a new approach to predict ground slope angles based on the return signal of full waveform LiDAR independently of any DEM. This approach addresses the issues sloped terrain has posed within the environmental remote sensing scientific community for some time now, as slopes increase the likelihood of spurious derivations of biophysical parameters such as vegetation height, crown size, and biomass estimates. Applying the model to spaceborne LiDAR waveforms shows great potential for accurate identification of ground slope.
The model relies on some initialisation data, consisting of known slope values that correspond to the identified waveform features used by the model in it's prediction process. This initialisation data acts as a basis from which slope predictions can be made. Using the model, slope predictions were made for a) simulated waveforms for slopes between 0° and 20° from the FLIGHT radiative transfer model, which considers forest geometries and parameters to map the path of incident photons from source to detector in order to build up a waveform image; and b) waveforms from the Geoscience Laser Altimeter System (GLAS), previously on board the Ice Cloud and land Elevation Satellite (ICESat). For GLAS, ground slope derived from airborne high-resolution discrete point return LiDAR was used as the initialisation data. We found that slope observations and predictions were highly correlated for both simulated waveforms and GLAS waveforms with a significant Pearson correlation coefficient of R^2=0.99 and R^2=0.89 respectively. Slope observations and predictions agree very well over low to mid-range slopes (0° -10°), in most cases, whereas predictions of more severe slopes (>18°) produce greater variability. Results also suggest that parameters other than slope, e.g. the presence of thick understorey, may influence the shape/size of any waveform's ground component. This will lead to an increased likelihood of spurious slope predictions being made. Future research will aim to further develop the model to allow for the correction of LiDAR based forest metric derivations with respect to slope and to allow application of the model to all LiDAR waveforms where initialisation data is available.

Alberta's Wet Areas Mapping Initiative

Barry White[1], Jae Ogilvie[2], Doug Hiltz[3], Kim Wen[4] and Paul Arp[5]

[1]Alberta Environment and Sustainable Resource Development
barry.white@gov.ab.ca

[2]Faculty of Forestry and Environmental Management, University of New Brunswick
jae.ogilvie@unb.ca

[3]Faculty of Forestry and Environmental Management, University of New Brunswick
c3acn@unb.ca

[4]Faculty of Forestry and Environmental Management, University of New Brunswick
kimhwen@gmail.com

[5]Faculty of Forestry and Environmental Management, University of New Brunswick
arp2@unb.ca

Paper Number: SL2012-028

Abstract

Sustainability of Alberta's sensitive aquatic habitats and forests is under risk due to unprecedented land use challenges. Innovative planning solutions that are economic, timely and ensure positive outcomes are sought by Albertans. Alberta has been working closely with researchers at the University of New Brunswick since 2004 to develop and test the effectiveness of a depth-to-water table mapping tool. This initiative, recently honoured for innovation excellence by the Alberta Science and Technology Foundation, is lead by Alberta Environment and Sustainable Resource Development. Scope spans some 22 million hectares including the province's foothills and boreal regions. Spatially explicit datasets predict the location of small water bodies such as intermittents, often as small as 10cm in width, and wet, saturated soils which are currently not known to resource planners but yet are sensitive to disturbance. Datasets also identify the locations of dry, drought prone areas where limited soil moisture may limit forest productivity. As a result of this longstanding research partnership, approximately 14.2 million hectares of lands have been mapped as of this date. Alberta's mapping process relies upon newly acquired discreet pulse, airborne LiDAR data to produce maps of superior quality with a resolution of 1m. Significant efforts are currently underway to explore opportunities for this new planning approach to influence policy and operational practices within the forestry and energy sectors. Efforts are also underway to explore how wet areas mapping in Alberta can inform complementary research in areas such as access management, silviculture, growth and yield, biodiversity, reclamation, recreational trail and road routing, and emergency spill management.

Using terrestrial LiDAR to assess structural diversity across multiple forest types in Australia

Kim Calders[1], Glenn Newnham[2], John Armston[3,4], Nicholas Goodwin[3], Darius Culvenor[2], Jasmine Muir[3], Melissa Fedrigo[5], Craig R. Nitschke[5], Martin Herold[1]

[1] Laboratory of Geo-Information Science and Remote Sensing; Wageningen University; Droevendaalsesteeg 3, Wageningen 6708 PB, The Netherlands; kim.calders@wur.nl

[2] CSIRO Land and Water; Private Bag 10, Clayton South, VIC 3169, Australia

[3] Remote Sensing Centre; Department of Science, Information Technology, Innovation and the Arts; Ecosciences Precinct, 41 Boggo Road, Dutton Park QLD, Australia, 4102

[4] Joint Remote Sensing Research Program; School of Geography, Planning and Environmental Management; University of Queensland; Brisbane, Australia, 4072

[5] Department of Forest and Ecosystem Science; School of Land and Environment; University of Melbourne; 500 Yarra Boulevard, Richmond, Melbourne, VIC 3121, Australia

Paper Number: SL2012-037

Abstract

The introduction of terrestrial LiDAR (Light Detection And Ranging) allows for measuring forest structure with high detail and accuracy. Terrestrial laser scanning could be used for the calibration in broad scale biomass mapping. Earlier work on the use of terrestrial LiDAR in Australian forests (Lovell et al., 2003) showed the potential of terrestrial LiDAR to measure gap probability and plant density profiles but the value of these parameters in biomass estimation has not been fully explored. It is still unclear if parameters inferred from LiDAR scans, such as canopy cover and tree height, are required for accurate biomass estimation or if there is a link to directly inferred parameters (e.g. gap probability and plant density profiles) from LiDAR scans.

Terrestrial LiDAR data has been acquired in different Australian forest types during the period March – June 2012. The forest types are: open woodland, Eucalypt open forest and sub-tropical rainforest. Data was collected in 14 plots, covering a wide range of plot biomass. Each plot consisted of 5 scan locations. The terrestrial LiDAR data has been acquired with the RIEGL VZ-400 3D terrestrial scanner. The scanner settings were the same for all the data collection and a similar plot setup was used across the different forest types. The RIEGL VZ-400 scanner collects multiple return LiDAR data (up to four returns per emitted pulse) as well as additional waveform data.

Here, we compare and contrast plant density profiles from terrestrial LiDAR data for different forest types in Australia. Metrics derived from these profiles are then used to estimate biomass. Reference biomass data from field measurements is available for all the forest types and is used to validate the biomass inferred from the terrestrial LiDAR data. The inferred biomass will also be compared to estimates derived from established allometric equations.

References

Lovell J.L., Jupp D.L.B., Culvenor D.S., Coops N.C., 2003. Using airborne and ground-based ranging lidar to measure canopy structure in australian forests. *Canadian Journal of Remote Sensing*, 29, 607-22.

Comparing ALS and TLS Eucalypt forest plot 3D voxel distributions under different sampling scenarios

Chris Hopkinson[1,2], Jenny Lovell[1], Laura Chasmer[2], Eva van Gorsel[1], Darius Culvenor[3], Natascha Kljun[4]

[1]Centre for Marine and Atmospheric Research, CSIRO, Canberra, Australia
Hopkinsoncd@gmail.com
[2]Cold Regions Research Centre, Wilfrid Laurier University, Waterloo, ON. Canada
[3]CSIRO, Land and Water, Melbourne, Australia
[4]School of the Environment & Society, University of Swansea, Swansea, UK

Paper Number: #SL2012-039

ALS and TLS data can be collected under a range of conditions and configurations. Uncertainties exist concerning the consistency and transferability of modeled 3D canopy attributes under different laser sampling regimes. Four point clouds (two ALS and two TLS) are compared for a 20m radius Eucalypt forest plot near Tumbarumba in the Bago State Forest, New South Wales. The first ALS dataset was collected in 2001 (ALS01) using an Optech ALTM (~ 1000 m agl, 25 kHz PRF, <40 deg FOV). Another ALS dataset was captured in 2009 (ALS09) using a Riegl LMS-Q560 (~ 300 m agl, 180 kHz PRF, 60 deg FOV). The two TLS datasets were collected in 2009 using an Optech ILRIS 3D frame scanner with 40 deg x 40 deg FOV and the Echidna Validation Instrument (EVI) hemispherical scanner. The ILRIS was ~ 1.5m above ground level and ~40m outside the plot to capture the full plot width, with two scans vertically stacked (upward scan of ~ 25 deg) to capture the canopy profile. The EVI was operated ~ 2m above ground at the plot centre. All scans were aligned and aggregated to 2m x 2m x 2m voxels to facilitate 3D comparisons. Sample density was highly variable, with ALS01, ALS09, ILRIS and EVI returning 1840, 38276, 412717 and 520229 returns throughout the plot. Due to overhead sampling geometry, ALS01 and ALS09 displayed uniform and similar spatial sampling with grid cell to grid cell standard deviations of 76% and 78% the mean sampling density, respectively. Given the lateral upward FOV of the ILRIS and EVI scans, spatial sampling standard deviations were 187% and 413% of the mean, respectively. Due to scan proximity, ILRIS sampling density reduced by ~1.2% / m through the plot, with voids being more prevalent on the distal side. For the EVI scan, 50% of the sampling occurred within a 6m x 6m horizontal region centred on the scanner, with density displaying exponential decay with distance. The variable sampling geometries led to significant differences in all vertical sample point density distributions. EVI peaked at ground level (due to short range), ILRIS displayed a broad peak in the centre of the canopy (due to side scan viewing), while both ALS datasets displayed peaks at canopy top and ground (growth and phenological changes between ALS datasets precludes more detailed comparison). Comparing vertical plot volume sampling distributions illustrated that all configurations captured peaks at canopy and ground level. However, relative proportions, heights and depths of the ground and canopy peaks illustrated large differences with total plot volume sampled for ALS01, ALS09, EVI and ILRIS being 14%, 40%, 31% and 38%, respectively. ALS09 and ILRIS canopy volume profiles were similar, suggesting that at this plot, occlusion from within-plot hemispheric sampling under represented upper canopy foliage more than the externally captured scans. While each scanning configuration is optimal for a given canopy sampling application, the marked differences between all distributions indicates that canopy attribute models are not easily transferable between point clouds captured under differing scan configurations and geometries unless systematic sampling biases are corrected.

Estimation of forest layer structure and canopy density using airborne laser scanning data

Werner Mücke and Markus Hollaus[1]

[1]Institute of Photogrammetry and Remote Sensing, Vienna University of Technology
[wm,mh]@ipf.tuwien.ac.at

Abstract Number: SL2012-043

Canopy density (CD) is described as the percentage of terrain area overgrown by tree crowns or vegetation, in general. It is an indicator for important bio-physical parameters, such as light transmittance or interception. Usually it does not account for the stratification of plants below the canopy. However, research has shown that structural diversity in forests, both in horizontal and vertical direction, is a key indicator for various forest functions. These include productivity, habitat quality and protection against natural hazards, such as landslides, avalanches or rock fall. Thus, acquisition of vegetation layer structure and CD is part of modern field inventory in forestry, although most often only in the form of random or regular sampling within the inventory plots.

Within this study a method for area-wide estimation of forest layer structure and canopy density is described. It follows the applicable definition in the Austrian forest inventory, which states a classification into three vegetation layers (canopy and two sub-canopies) at fixed height intervals. The study site is located in the *Stand Montafon* of the federal state of Vorarlberg in the western part of Austria and is dominated by coniferous tree species. A medium density airborne laser scanning (ALS) point cloud (~4 echoes per m²) acquired under leaf-off conditions is used for the analysis. The proposed methodology exploits the fact that ALS can pass through small gaps in the foliage or intertwined branches and can directly map sub-canopy vegetation in 3D. The vertical distribution of the laser echoes allows drawing conclusions on the structural complexity of the vegetation and, moreover, the occurrence of multiple sub-canopy layers overgrowing each other. Additionally, it has to be assumed that occlusion by upper layers reduces the probability for lower ones to be reached by the laser pulses. A point cloud based, but nevertheless computationally fast approach for the characterization of the canopy and sub-canopy layers is proposed. The method accounts for the above mentioned occlusion effect by considering the dependency between overlapping layers in a multiple linear regression model for the prediction of the vegetation layer coverage. Raster maps describing the coverage for each of the three canopy layers are created. CD is subsequently derived on a raster basis as the aggregation of the derived canopy layer maps.

For training and evaluation purposes two data sets, namely 488 inventory plots and a more detailed 500 x 50 m transect of manually collected forestry data including crown radii and vegetation layer coverage, are used. Results show that the method works reliably for the top most layer ($R^2 = 0,82$) and CD ($R^2 = 0,72$), whereas it tends to underestimate the coverage of the two sub-canopies ($R^2 = 0,68$ and $0,53$). This drawback can be overcome using full-waveform ALS data, which tends to produce more echoes in general and depicts lower layers with higher detail. The presented method incorporates operative forest inventory data and district-wide ALS-data, thus it is operationally applicable for area-wide estimation of forest layer structure.

Does forest structure affect the Normalized Difference Vegetation Index?

Naoko Miura

Graduate School of Agricultural and Life Sciences, the University of Tokyo,
miura@uf.a.u-tokyo.ac.jp

Paper Number: SL2012-045

Abstract

An efficient and cost-effective total sampling assessment tool is required for better natural resource management. Remote sensing data derived from satellite and airborne sensors is superior to field survey data in terms of high-spatial coverage, near simultaneous acquisition, repeated regional accounting and cost. To date, most natural resource remote sensing has been undertaken using passive sensing technologies, mainly in the visible/shortwave infrared red portions of the electro-magnetic spectrum. The Normalised Difference Vegetation Index (NDVI) computed from these portions of the electro-magnetic spectrum is one of the most widely used techniques to make quantitative estimates of vegetation properties. In forestry, numerous papers have documented the utility of LiDAR for the estimation of forest attributes. The synergy between those techniques might extend application potential greatly. However, it has been reported that the NDVI values are susceptible to various physical effects such as the atmospheric condition at the time of data acquisition, topography and bare soil under the vegetation cover. Furthermore, when the NDVI value is calculated for forested area, it is also the concern what structural components reflect mostly on the value. Therefore, in order to obtain better understanding of the NDVI value in the forested landscape, the relationship between forest structure and the NDVI values derived from SPOT-5 satellite image in Eucalyptus forest was examined using LiDAR derived forest structure variables: opening above the ground, low vegetation and medium vegetation, canopy cover, the amount of low, medium and high vegetation, and vertical density of high trees. Preliminary statistical analysis with Pearson correlation coefficient showed that the mean NDVI value was positively correlated with the amount of low and medium vegetation. This indicates that the presence/amount of mid- and understorey vegetation under canopy trees may also have an impact on determining the NDVI values in forested landscape.

Radiometric Analysis of Mulit-Wavelength Airborne Laser Scanning Data of Different Case Study

Christian Briese[1], Martin Pfennigbauer[2], Andreas Ullrich[3] & Norbert Pfeifer[4]

[1]Institut of Photogrammetry and Remote Sensing of the Vienna University of Technology, Vienna, Austria, and LBI for Archaeological Prospection and Virtual Archaeology, Vienna, Austria, cb@ipf.tuwien.ac.at
[2]*RIEGL* Laser Measurement Systems GmbH, Austria, mpfennigbauer@riegl.com
[2]*RIEGL* Laser Measurement Systems GmbH, Austria, aullrich@riegl.com
[4]Institut of Photogrammetry and Remote Sensing of the Vienna University of Technology, Austria, np@ipf.tuwien.ac.at

Paper Number: #SL2012-047

Abstract

Airborne laser scanning (ALS) is widely used for the acquisition of urban and rural areas. Nowadays a wide range of ALS sensors with different technical specifications can be found. One key parameter is the laser wavelength which determines the instrument's relative ranging capabilities due to the wavelength-dependent backscatter characteristic of the sensed surfaces. Next to the geometric information (i.e. the location in space) current ALS sensors usually record amplitude information for each echo. In order to utilize this information for the study of the backscatter characteristics of the sensed surface, radiometric calibration is essential. This paper focuses on the radiometric calibration of multi-wavelength ALS data as opposed to conventional monochromatic (single-wavelength) ALS data. After a short introduction theory and practice of the radiometric calibration of ALS data based on in-situ reference surfaces is presented. Based on individual monochromatic radiometric reflectance readings a calibrated multi-wavelength reflectance image can be generated. It is important to note that this image is not influenced by shadows (due to the active illumination of the sensed surface) and from a geometric viewpoint the position of the objects on top of the terrain surface is not altered (true orthophoto). Within this paper the approach is demonstrated for two projects both comprising three different single-wavelength ALS data acquisitions (532nm, 1064nm, and 1550nm). One project covers the area of the city of Horn in Austria and the other covers a rural area near Horn containing woods, fields, open grass land, a small lake, and a small village.
The practical results presented here show the applicability of calibrating multi-wavelength radiometric imagery from ALS data and provide an insight into the challenges of radiometric processing and exploitation of multi-wavelength ALS data. Based on the resulting multi-wavelength reflectance information spectral analysis of the radiometric behavior of the sensed surfaces at the three different wavelengths is possible. Furthermore, the analysis of the ALS data focuses on the point clouds obtained with the three different laser scanners with respect to point density, multi-target resolution, scan alignment, vegetation penetration, and water surface and ground caption.

Dual-wavelength laser scanning for measurement of vegetation moisture content

R. Gaulton[1], F. M. Danson[2] & F.A. Ramirez[2]

[1] School of Civil Engineering & Geosciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK, rachel.gaulton@ncl.ac.uk
[2] School of Environment and Life Sciences, University of Salford, Salford, M5 4WT, UK, f.m.danson@salford.ac.uk

Paper Number: SL2012-051

Abstract

Biochemical properties of vegetation, such as moisture content, can be important indicators of plant health, plant physiological response to environmental stress and of disease. These are issues of global concern with increased frequency and severity of droughts in many forested regions and rising rates of spread of forest pests. Dual-wavelength laser scanning provides a promising approach to detect early signs of vegetation stress with significant advantages over passive remote sensing approaches. Firstly, laser scanning is not reliant on illumination conditions and secondly, it allows separation of the influence of canopy components and of the understorey or soil on estimates of biochemical parameters.

The Salford Advanced Laser Canopy Analyser (SALCA) is a unique hemispherically-scanning, full waveform terrestrial laser scanner that operates using two laser wavelengths (1064 nm and 1545 nm), theoretically allowing calculation of ratios sensitive to vegetation moisture content. This paper presents initial results of leaf optical properties modelling and leaf-level laboratory measurements to assess this capability and determine sensitivity to moisture content. Through repeated scans of drying leaves, a strong almost linear relationship (R2=0.80) was found between a normalised ratio of the 1545 nm and 1064 nm wavelengths and Equivalent Water Thickness (EWT) of the leaves. However, significant within-leaf variation in the ratio was also detected, most likely due to stomatal heterogeneity in response to moisture stress, resulting in fine-scale variations in EWT within the beam footprint.

The results suggest potential for use of the SALCA instrument to up-scale physiological measurements to canopy scales, allowing improved validation of airborne and satellite measurements, to measure spatial heterogeneity in canopy moisture stress and to provide an indication of the potential of future airborne dual or multiple wavelength LiDAR sensors for measurement of leaf biochemical parameters. Research is on-going to determine the ability of the SALCA instrument to estimate tree health at canopy scales.

Finding snow in the mountains: Using LiDAR to quantify snow volume in heterogeneous landscapes

Wade T. Tinkham[1], Alistair M.S. Smith[1], Tim E. Link[1], Michael J. Falkowski[2], & Hans-Peter Marshall[3]

[1]Forest Rangeland, and Fires Sciences, University of Idaho; tink8633@vandals.udiaho.edu, alistair@udiaho.edu, tlink@udiaho.edu
[2]School of Forest Resources and Environmental Science, Michigan Technological University; mjfalkow@mtu.edu
[3]Department of Geosciences, Boise State University; hpmarshall@boisestate.edu

Paper Number: SL2012-052

Abstract

There is increasing need to quantify snow and its uncertainties in complex terrain using remote sensing, especially in water-limited regions sensitive to climate change. Here we develop a novel method to examine the accuracy of snow volume estimates across seven land cover types in a semi-arid mountainous catchment by coupling intensive topologic surveying, machine learning, multi-temporal LiDAR, and coincident manual snow observations. Results show that LiDAR can be used to accurately quantify the distribution of snow depth across cover types ranging from bare ground to mature forests. Snow drifts, which are important for maintaining spring and summer streamflow and establishing and sustaining water-limited plant species, contained 30 ± 6% of the snow volume while only occupying 10% of the catchment area. Quantifiable accuracies of catchment-scale snow volume estimates not only illustrate the utility of multi-temporal LiDAR for snow studies but will also reduce uncertainty in seasonal water availability estimates.

Modelling the internal canopy radiation regime using terrestrial laser scanning data

Van Leeuwen, M.[1], Coops, N.C.[1], Hilker, T.[2], Wulder, M.A.[3], Newnham, G.[4] & Culvenor D.[4]

[1] Forest Resources Management, University of British Columbia, 2424 Main Mall, Vancouver, V6T 1Z4, BC, Canada, mvanleeu@interchange.ubc.ca, nicholas.coops@ubc.ca

[2] NASA Goddard Space Flight Centre, Biospheric Sciences Branch Code 618, Bld 33, #G310, 8800 Greenbelt Road, Greenbelt, MD 20771, USA, thomas.hilker@nasa.gov

[3] Pacific Forestry Centre, 506 Burnside Road West, Victoria, British Columbia, V8Z 1M5, BC, Canada, Mike.Wulder@NRCan-RNCan.gc.ca

[4] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Locked Bag 10, Clayton South VIC 3139, Australia, Glenn.Newnham@csiro.au, Darius.Culvenor@csiro.au

Paper Number: #SL2012-054

Abstract

(to be considered as part of the Dr. Alan Strahler terrestrial laser scanning community):

Modelling the internal canopy radiation regime has provided significant insight into the productivity and functioning of forest ecosystems. Vertical profiles of photochemically active radiation affect the allocation of nutrients and efficiency of photosynthesis. The vertical distribution of photochemically active radiation is primarily affected by canopy structure that encompasses the distribution of foliage elements, stocking density and crown geometries. Radiative transfer models may be used to investigate how canopy structure influences the distribution of PAR with canopy depth. Through integration with physiological models, these radiative transfer models facilitate studies into the effects of canopy structure on primary productivity and resilience against disturbance, such as drought events for example.

In remote sensing, radiation transfer has primarily been modelled over broad scales, resulting in spatial and temporal averages of canopy radiation. Fusing of spatially and temporally averaged canopy radiation with physiological models may lead to uncertainty, however, as a leave's ability to adapt to instantaneous irradiation conditions is different from its ability to adapt to prevalent irradiation. Modelling at spatially and temporally finer scales is challenging due to the cost and labour needed to acquire canopy structural information, as well as processing needed to obtain canopy radiation from data about canopy structure.

This poster presents on results obtained from an automated pipeline for geometric explicit modelling of canopy structure from point cloud information. The geometric model is used to simulate the fraction of absorbed photochemically active radiation at arbitrary locations within the canopy. Results show consistent behaviour between simulated vertical profiles of fPAR, and canopy gap probabilities derived from full waveform data. The use of these models for coupling shoot-level physiology to changes in instantaneous irradiation is discussed.

Solar-Laser double light source vegetation canopy model: construction and validation

Gao shuai[1], Niu Zheng[2], Wang Lijuan[3], Wang Li[4], Wu Mingquan[5], Xu Shiguang[6], Hao Pengyu[7], Wang Xinlei[8]

The State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China; and
[1]gaoshuai@irsa.ac.cn,[2]niuz@irsa.ac.cn, [3]muzixin8866@163.com, [4]wangli@irsa.ac.cn, [5]wumq@irsa.ac.cn, [6]haibinggis@163.com, [7]haopy8296@163.com, [8]xinjm0508@sina.com

Paper Number: #SL2012-057

Abstract

When the vegetation was illuminated by laser in daytime, the sensor usually received two types of light intensity in the spectrum range: sun and laser itself. In this situation, the conventional bidirectional intensity distribution would vary with two incidences light. Theoretically, this would supply more canopy information if the reflection principle was carefully clarified. So, in this paper, Solar-Laser Double Source Canopy Model (DSCM) were constructed and analysed. Firstly, the Li-Strahler Geometric Optical (GO) model was modified by adding another input light source to simulate the intensity distribution. Then, the Radiative Transfer (RT) model was incorporated to express the canopy reflectance change with different incidence angle. So, the GO-RT mixed model was constructed with two incidents light. Based on the model, the reflectance phenomenon was simulated refer to the forest in the monochromatic light. It revealed that, by adding another incident light (laser), two "hotspot" was constructed and that was influenced by both incident light intensity and incident angle, also it was affected by the forest canopy structure. After the simulation, the experiment was carried out for the ideal forest condition constructed with plastic ball illuminated by sun and laser (808nm) in daytime. Then in night, the experiment was repeated. With the variation of light source and "forest" destiny, bidirectional reflectance distribution function (BRDF) of sun, laser and double source were observed respectively. And the experiment results were basically accordance to the DSCM simulation, especially for the hotspot phenomenon. So, it was convinced that DSCM could be applied for the vegetation biophysical and biochemical parameters retrieval study with multi-spectral laser became mature.

References

ROSS, J., 1981, The radiation regime and architecture of plant stands: Springer.

LI, X., STRAHLER, A. H., and WOODCOCK, C. E., 1995, A hybrid geometric optical-radiative transfer approach for modeling albedo and directional reflectance of discontinuous canopies. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 466-480.

KUUSK, A., 2001, A two-layer canopy reflectance model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 71, 1-9.

On the boresight LiDAR estimation

Gil Gonçalves[1,2] & André Jalobeanu[3]

[1]INESC-Coimbra and [2]University of Coimbra, Portugal, gil@mat.uc.pt
[3]Geophysics Centre of Évora,University of Évora, Portugal, jalobeanu@uevora.pt

Paper Number: 058

Abstract

Within the AutoProbaDTM project (sites.google.com/site/autoprobadtm/), we plan to develop fast and fully automated techniques to derive topographic maps and error maps, from full-waveform airborne LiDAR data. A probabilistic approach is used in order to modelling surfaces and data acquisition, solving inverse problems and handling uncertainty. Bayesian inference provides a rigorous framework for unsupervised reconstruction of the DEM and error propagation from the data to the end result, treating all quantities as random variables. Automatic sensor calibration plays a major role in this project. In fact, the overall positional accuracy and uncertainty obtained from the LiDAR technology depends on the assembly and calibration of the three system components: the GPS (Global Positioning System), the IMU (Inertial Measurement Unity) and the laser-scanner device. In this poster we propose a novel method based on the Bayesian inference to address this problem as well. Because this method does not use planar surfaces it is well suitable for forest areas with poor geometric planar features such as building roofs. The first contribution is to use not only the 3D points extracted from the raw waveforms but their uncertainty as well, and to apply a probabilistic surface matching with spatially variable point accuracy in order to obtain the attitude corrections. The second contribution consists of using all the flight lines, where most methods only use the calibration cross. In this way we can also estimate the attitude drift of the sensor platform and correct the LiDAR data for temporal attitude variations.

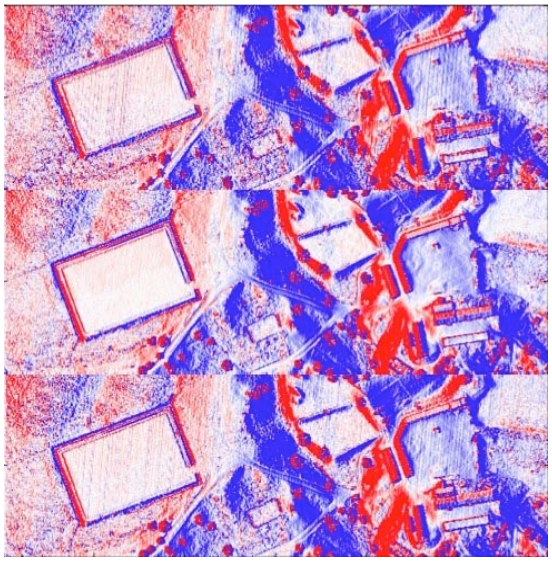


Figure 1: Hi-frequency IMU errors

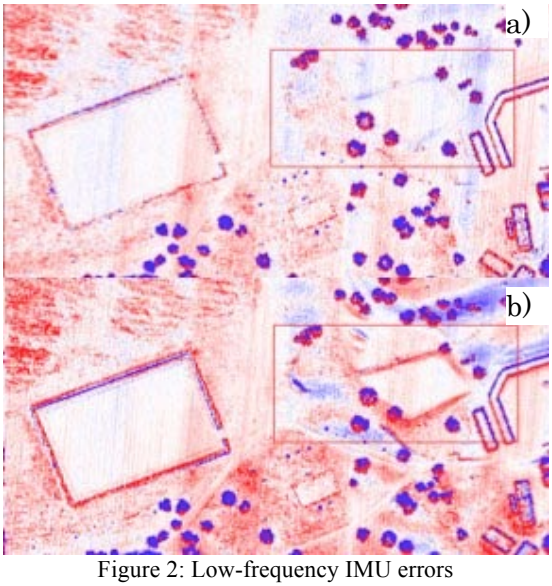


Figure 2: Low-frequency IMU errors

Figure 1 shows the corduroy effect (high-frequency IMU errors), visible as stripes on slopes parallel to flight lines, for 3 overlapping strips with different scan angles. Figure 2 shows the low-frequency IMU errors (due to temporal correlation) after (a) and before (b) correction by optimal shifting on a small area.

Finally, we use the probabilistic framework for error propagation and we propose a probability distribution of the estimated boresight angles.

Operational Forest Management using LiDAR derived Tactical Net Harvest Area in NSW Native Forests.

Tony Brown[1]

[1]Forests NSW, Australia tony.brown@sf.nsw.gov.au

Paper Number: SL2012-065

Abstract

Traditional strategic systematic random sampling estimates of product volume have proven accurate at the estate level, however provide little useable volume or product information at forest or compartment level, a requisite for valid tactical and operational planning. In mid 2011 FNSW native forest regions were handed a management challenge to produce a fully ground verified forest harvesting schedule with reduced staff numbers and diminishing preferred species across a resource limited estate. FNSW has significant areas of the native forest estate on the north coast captured with LiDAR, and had developed enough skill and confidence in LiDAR use to look to this technology as a significant contributor in meeting this challenge.

Traditional vegetation information such as forest typing has been derived from API over many years by many operators of varying experience and skill levels. Resolution and accuracy of forest typing vary significantly, while forest disturbance and harvesting further compromise data accuracy. LiDAR derived canopy height model investigation led to a structural stratification of the native forest. Through integration of mean CHM and a count of 'tall' pixels within a given focal window, the native forest has been classified into 5 structurally discrete strata. Whilst structure is not entirely related to product volume, its correlation does provide significant opportunity to refine and improve product volume estimates at small scales, which can then be sampled on ground by experienced staff.

Existing regulation around stream networks require FNSW to manage streams as mapped on 1:50,000 topographic maps (derived from API anywhere up to 40 years previous) differently to those which are unmapped. FNSW investigated correlations between mapped streams and those derived through use of LiDAR data, and by analysing stream length by regulatory class an approximation using catchment area was derived. These 'contributing' streams were buffered as a grid using corresponding regulatory constraints, and integrated with existing mapped reserves such as old growth or owl habitat.

Ground based harvesting operations are limited under regulations to operating on slopes <30 degrees, while in practice very little volume is realised off slopes >25 degrees. Regulation around slope relates to infield identification, leading to a realistic interpretation of LiDAR derived slope to accurately reflect how slope is managed during a forest operation.

Slope, strata and reserve datasets were converted to 5m grid cell, combined within ArcGIS providing a single raster output with 110 classes (5 strata, 11 slope and 2 reserve) which has become known as the 'Tactical Net Harvest Area' (NHA). This Tactical NHA provides an extremely powerful planning tool for forest management, which while not conforming precisely with regulations, FNSW claim significantly better represents the area available for harvest access at sub compartment scale over existing spatial data.

Using LiDAR derived products, FNSW met a significant management challenge in preparation of a fully ground verified operational plan. Highly spatially accurate LiDAR vegetation data provides a level of resource information previously unavailable, while LiDAR derived streams have now prompted interagency regulatory review, likely to result in a more cost effective and environmentally focused regulatory environment.

The detection and characterization of broad-leaved forest canopy gaps: a regeneration perspective

Stéphanie Bonnet[1], Sébastien Bauwens[1], François Lehaire[1], Gauthier Ligot[1] & Philippe Lejeune[1]

[1]Univ. Liège, Gembloux Agro-Bio Tech, Unit of Forest and Nature Management, 5030 Gembloux, Belgium, s.bonnet@ulg.ac.be, sebastien.bauwens@ulg.ac.be, francois.lehaire@ulg.ac.be, gligot@ulg.ac.be, p.lejeune@ulg.ac.be

Paper Number: sl2012-069

Abstract

Canopy gaps are areas of high regeneration potential and, in uneven-aged forest, gaps are therefore places of particular care for the forest manager. Nevertheless the cartography and characterization of canopy gaps are complex issues. This paper addresses the fundamental question of the canopy gap definition: what is the minimal area, the maximal height of vegetation, type of regeneration, etc? From a regeneration point of view, canopy gaps can be defined as holes in the forest cover where light conditions are suitable for recruitment. As an active sensor, LiDAR has made it possible to tackle the problems of shadows and penetration into the canopy, typical of aerial images. This study investigates the cartography and characterization of forest canopy gaps as areas of natural regeneration.

Remote sensing data are composed of leaf-off and leaf-on LiDAR high density datasets (13 points per m²) with intensity information and a point classification. Very high resolution (25cm) multispectral (r,g,b and ir) aerial images were also captured for the summer dataset. The methodology follows a number of steps: the detection of the canopy gap, the correction of this first cartography and the resultant characterization. Results are compared between leaf-off (winter) and leaf-on (summer) conditions.

The detection of gaps is based not only on height criteria but also on "canopy porosity" metrics (e.g. the penetration rate of pulses to the ground). Thresholding and OBIA approaches are then used from these height and porosity metrics. Following this first cartography, a filtering and cleaning limits step (based on several criteria including dimension and shape) was crucial. The canopy gap description comprises two parts; first, the gap is described with dimensions, shape and vegetation cover information. Secondly, the structure and stratification of the gap and the surrounding forest is characterized within an extended area delineated around the gap. This "extended" forest canopy gap can also be described with ancillary data extracted from LiDAR, image or vector database, including altitude, slope, aspect, soil, etc.

In conclusion, the development of a "regeneration potential index" based on the delineation and characterization of the canopy gaps would be extremely useful from a forest planning perspective. Mapping gaps would also enable analysis at a landscape level with regard to the stand scale successful regeneration, biodiversity and wildlife management. Furthermore, the exploitation of hemispherical photographs would be an interesting complementary issue to link regeneration and light needs.

Comparison of allometric and TLS-feature based models in crown biomass determination

Ville Kankare[1], Markus Holopainen[1], Mikko Vastaranta[1], Eetu Puttonen[2], Xiaowei Yu[2], Matti Vaaja[2], Juha Hyyppä[2], Hannu Hyyppä[3,5] & Petteri Alho[4,3]

[1]University of Helsinki, Department of Forest Sciences, Finland, first.last@helsinki.fi
[2]Finnish Geodetic Institute, Finland, first.last@fgi.fi
[3]Aalto University, Research Institute of Measuring and Modelling for the Built Environment, Finland, first.last@aalto.fi
[4]University of Turku, Department of Geography and Geology, Finland, first.last@utu.fi
[5]Helsinki Metropolia University of Applied Sciences, Finland, first.last@metropolia.fi

Poster Number: #SL2012-070

Abstract

Forest biomass quantity and its changes are difficult and time-consuming to measure accurately. Accurate biomass mapping has numerous applications in cases of forest-bound carbon, bioenergy, -fuel and forest hazard monitoring. A major proportion of the total forest carbon storage consists of the growing stock's carbon reserves. Determination of individual tree-level (ITL) biomass requires accurate measurements of tree trunk, branches and needles. Accurate field measurements can be done only in a destructive manner and therefore non-destructive and more cost-efficient methods are needed. Fixed-position (mounted on a tripod) terrestrial laser scanners offer a great potential for three-dimensional (3D) mapping of smaller areas with high detail. Terrestrial laser scanning (TLS) is an efficient and objective option for acquiring accurate field data. TLS measurements provide dense point clouds in which features describing biomass can be extracted from trunk and canopy dimensions. In Finland biomass models have been developed but the models do not fully utilize canopy size information. Models are mostly based on easily measurable parameters on field such as tree species, diameter-at-breast-height (DBH), height and crown length. We hypothesized that TLS measurements would increase the estimation accuracy of ITL crown biomass. The main objective of this study was to estimate ITL crown biomass, based on models developed using TLS data, for Scots pine and Norway spruce. Describing features were extracted from individual tree point clouds and used in biomass model development. The modelling dataset included 64 laboratory-measured trees. Models were developed for living branch and dead branch biomass and the results were compared with existing biomass models.

The results showed that crown biomass estimation accuracies were improved, compared with those of existing models. Root mean square errors (RMSEs) for living branch biomass were 23.36% for pine and 38.13% for spruce using models developed from TLS. Relative RMSE for dead branches was higher because the amount of dead branches is low and good describing features are difficult to find. Development of robust TLS-based biomass models is challenging, because it would require large treewise laboratory-measured reference datasets and the laser features used as explanatory variables should be nondependent on the scanning parameters. We suggest the use of automatically extracted crown size measurements from TLS as a basis for allometric crown biomass models. In the current state, operational TLS measurements are not fully applicable under various forest conditions but if TLS processing becomes more automatic and reference data could be collected even with the use of mobile platforms, it could also provide savings in reduced fieldwork. However, feasible measuring of crown and stem characteristics, that are laborious to measure by traditional means accurately, is the strength of TLS.

Diagnosing structure and composition typologies in uneven-aged broad-leaved forests: a comparison of classification methods

Stéphanie Bonnet[1], Yves Brostaux[2], Hugues Claessens[1], Philippe Lejeune[1]

[1] Univ. Liège, Gembloux Agro-Bio Tech, Forest and Nature Management, 5030 Gembloux, Belgium – s.bonnet@ulg.ac.be, hugues.claessens@ulg.ac.be, p.lejeune@ulg.ac.be

[2] Univ. Liège, Gembloux Agro-Bio Tech, Applied Statistics, Computer Science and Mathematics, 5030 Gembloux, Belgium - y.brostaux@ulg.ac.be

Paper Number: sl2012-071

Abstract

Structure and composition of forest stands are crucial factors for forest planning and biodiversity management. In Belgium, typologies of structure and composition exist to support planning in uneven-aged broadleaved forests (typically dominated by oak and beech). The principle of these typologies is to classify irregular stands with the percentage of small, medium, large, and very large trees (regarding dbh), and the percentage of basal area of oak and beech.
This paper investigates the potential of LiDAR data processed with classification methods (k-nn, K-Means, CART, etc.) to allocate a forest structure and composition type. For this purpose several supervised and unsupervised classification methods are compared, as well as the impact of leaf-on (summer) and leaf-off (winter) data to discriminate the forest types.
The study dataset is composed of a leaf-off and a leaf-on LiDAR high density data (13 points per m²). These datasets are characterized by intensity as well as a point classification (ground, high vegetation, water, etc.). Moreover, very high resolution (25cm) multispectral (r, g, b and ir) aerial images were also captured for the summer dataset.
Each field reference plot is characterized by several "LiDAR variables" calculated from metrics. These metrics are derived from the point cloud in raster format. Spectral information from aerial images is also integrated to complement LiDAR, particularly for species discrimination.
Reference data were obtained from field inventory through visual evaluation of tree girth based on a systematic sampling; this type of inventory is a very useful tool for the description of mixed broadleaved uneven-aged stands. A network of about 500 plots was visited in February 2012.

Using digital terrain models derived from airborne laser scanner data for detection of cultural remains

Ole Martin Bollandsås[1], Ole Risbøl[2], Terje Gobakken[1], Erik Næsset[1] & Anneli Nesbakken[2]

[1]Norwegian University of Life Sciences, Department of Ecology and Natural Resource Management, P.O. Box 5003, NO-1432 Ås, Norway. olebo@umb.no
[2]Norwegian Institute for Cultural Heritage Research P.O. Box 736 Sentrum, N-0105 Oslo, Norway. ole.risbol@niku.no

Paper Number: SL2012-072

Abstract

In Norway, all cultural remains predating 1537 are protected by the Cultural Heritage Act. Many of these cultural remains are found in forested areas and forest owners are obliged to make sure that they are not harmed during mechanized forest operations. Thus, in order to comply with the law, forest owners need to have accurate information on the locations of the different law-protected objects. Traditionally, cultural remains have been registered by means of field inventories which are expensive to carry out. For that reason areas have to be prioritized and focus has been on surveying areas outside forests. Studies have indicated that only about 10 - 20% of all cultural remains in forested areas are known. However, recent trials with manual interpretation on digital terrain models (DTMs) derived from airborne laser scanner (ALS) data have shown promising results on detection of cultural remains. Manmade anomalies will in most cases stand out compared to natural features because they tend to be somewhat regular in shape or that several anomalies form some kind of recognizable pattern. The results obtained so far indicate that the ALS data collected for regular forest inventory also can be exploited for geo-location and classification of cultural remains. The detection success, however, varies with several factors. The quality or resolution of the DTM is dependent on the point density of the ALS data and recent studies have shown that detection success is significantly affected by point density. The properties of the cultural remain in terms of size and shape is probably another factor, but no studies have addressed this question yet. Thus, the objective of the current study was to analyze if there were differences in detection success between cultural remains of varying sizes and shapes. Detection success was measured by the omission error. The study took advantage of data obtained from an experiment where four experienced archaeologists interpreted 81 DTMs each covering an area of 500x500 meter. The 81 DTMs originated from nine different blocks of land just north of Oslo, Norway. Thus, for each block, nine different DTMs were derived from ALS data with different properties in terms of point density (1, 5, and 10 p m^{-2}) and DTM-smoothing (three levels). The archaeologists interpreted the 81 DTMs in an individually randomized order according to a complete randomized block design. An analysis of variance (ANOVA) showed that there were significant differences between the omission errors of different classes of size and shape. There was a decreasing trend of the omission error with increasing size. The results also showed that circular and oval shaped remains were less frequently omitted compared to other shapes.

A case study using field surveys and LiDAR to quantify aboveground carbon, bird diversity, and tree species richness to prioritize conservation based on multiple ecosystem services

Samuel Finch[1], Lee A. Vierling[1], Kerri T. Vierling[1] & Andrew T. Hudak[2]
[1]University of Idaho, finc7201@vandals.uidaho.edu, leev@uidaho.edu, kerriv@uidaho.edu
[2]Rocky Mountain Research Station, ahudak@fs.fed.us

Paper Number: SL2012-174

Abstract

Quantifying and valuing ecosystem services at multiple spatial scales is central to understanding and managing the sustainability of global natural resources. The majority of attempts to quantify and value ecosystem services have focused on a single service, used proxy measures of services versus primary data, and/or failed to adequately integrate ecological and economic components into the research agenda. We integrated detailed field observations with LiDAR remote sensing and statistical modeling techniques to quantify variation in above-ground carbon, bird species richness, and tree species in a north Idaho mixed conifer forested landscape of ~20,000 ha. Additionally, all three ecosystem services were combined to identify ecosystem service "hotspots". Here, we examine the effect of spatial scale on the ability to quantify these ecosystem attributes, and present a case study that might be used to prioritize land conservation decisions. Scales of analyses ranged from grain sizes of 100m to 1500m. Values of carbon (Mg/ha), bird species richness (number of species standardized per 0.04 ha), and tree species variety (different species per 0.04 ha) were each given a standardized index value between 0 (minimum scaled value) and 1 (maximum scaled value). To quantify multiple ecosystem services hotspots, we summed the index value of the three individual ecosystem services, which resulted in an index value range of 0 (low) to 3 (high). Forests containing above-ground carbon values in the top 70% of the observed range were lost when grain size increased from 100m to 1500m. However, 84% of the total above-ground carbon occurred within the bottom three classifications (i.e. bottom 30%). Thus, while carbon hotspots were overlooked at larger grain sizes, the bulk of carbon is still accounted for at relatively coarse (1500m) grain size. Similar to the carbon dataset, bird species richness hotspot areas represented a small fraction of the total area. The concentration of species richness values, similar to carbon, resulted in bird species richness hotspots disappearing as grain size increases. Unlike carbon, both high and low bird species richness values were rare. Trends in tree species diversity as grain size increased were subtle because the top two diversity values (15.2% of total area at 100m grain size) disappeared when grain size increased from 100 m to 1500m. The combined ecosystem services dataset, which overlaid each individual ecosystem service, exhibited a 24% decrease in maximum index value (Index value of 2.03 at 100m grain size vs. 1.54 at 1500m grain size) when grain size increased from 100m to 1500m. Additionally, index values in the largest classification at 100m grain size represented 35% of the total area, but at the 1500m grain size, the area in the same classification increased to 50%. Therefore, the fine scale information afforded by LiDAR data may result in improved targeting schemes for prioritizing land conservation or payment for ecosystem services. Conversely, inaccurate and/or spatially vague information may propagate through the decision making and policy stages of ecosystem service conservation. Knowledge of the potential error and ways to mitigate scale related ecosystem service identification errors may therefore be useful in creating effective policy.

Estimation of Soil Surface Roughness from Airborne LiDAR: An Exploratory Study for SAR-based Modelling of Forest Biomass and Soil Moisture

Rocco Panciera[1], Russell Turner[2], Mihai Tanase[1] & Jorg Hacker[3], Ying Gao[4]
[1]CRC for Spatial information, The University of Melbourne, Australia
panr@unimelb.edu.au
[2]Remote Census, NSW, Australia
[2]Airborne Research Australia, School of Environment, Flinders University, Australia
[3]Department of Civil Engineering, Monash University, Melbourne, Australia

Paper Number: SL2012-075

Abstract

Soil surface roughness in the 10^{-2} – 10^{1} meter scale is the main factor determining surface backscatter from areas characterized by bare to low vegetated conditions. The parameterization of surface roughness and characterization of its spatial variability can pose major problems in forest biomass retrieval from Synthetic Aperture Radar (SAR) data since the total backscatter observed by the SAR sensors in areas of low vegetation density will include a significant direct backscatter from the ground which needs to be accurately accounted for in SAR retrieval models. This paper investigates the retrieval of surface roughness from airborne LiDAR observations as a mean to provide spatial surface roughness estimates for SAR applications.

Surface roughness is generally characterized through ground measurements of height variations along short (1-3m) transects. That information is mainly input into backscatter models using the heights Root Mean Square (RMS), correlation length and autocorrelation function, under the assumption that surface roughness can be represented by a single-scale stationary process. However, extensive research concerning surface roughness characterization has suggested that roughness parameters often show little or no spatial dependency, meaning that parameters measured at one position often poorly represent their surrounding. Moreover, scale dependency of the roughness parameters makes the assumption of a single-scale process unsuitable to accurately account for the complex geometry of natural surfaces, providing an additional source of uncertainty in SAR backscatter models. In addition, manual soil profiles are time consuming and labour intensive, hence there is a need to develop a faster and more cost-effective sampling alternative. LiDAR observations present an unprecedented opportunity to efficiently characterize surface roughness spatial distribution by providing high density, high precision soil surface heights measurements over large areas. A recent study has analysed the viability of such technique suggesting that the vertical height noise resulting from the LiDAR system noise is small enough to allow distinguishing cm-scale roughness variations.

This paper demonstrates for the first time the accuracy of estimates of surface roughness derived from airborne LiDAR observations over bare surfaces. Two airborne LiDAR surveys were conducted on September 5 and 22, 2011 over 15 bare agricultural paddocks in an area of intensive irrigated cropping in the Murrumbidgee catchment, Australia. Airborne data were supported by manual ground measurement of surface heights collected along 3 m profiles at 3 locations on each paddock. The accuracy of LiDAR-derived surface roughness estimates was assessed by addressing: (i) the repeatability of estimates of heights RMS, both between adjacent swaths and different dates, (ii) The error and correlation between LIDAR and ground estimates of heights RMS. Results showed that the LiDAR estimates were very stable between overpasses (-0.1±0.6cm mean difference). Excellent correlation was also observed between LiDAR and ground profiles ($r=0.85$, $p<0.05$), with a root mean square error of 0.88cm (+0.7cm mean error).

As an example application, the validated LiDAR estimates were then used to investigate surface roughness temporal changes due to weathering and tillage.

Modelling harvesting productivity with LiDAR

Alam, M. M.[1], Strandgard, M.[2], Brown, M.[3] and Fox, J. C.[4]

[1] The University of Melbourne, Australia, Email: mmalam@student.unimelb.edu.au
[2] University of the Sunshine Coast, Australia, Email: mstrandg@usc.edu.au
[3] University of the Sunshine Coast, Australia, Email: mbrown2@usc.edu.au
[4] The University of Melbourne, Australia, Email: jcfox@unimelb.edu.au

Paper Number: #SL2012-077

Abstract

Major factors impacting productivity and efficiency of a mechanised harvesting system include forest stand characteristics (tree size or piece size, stand density, undergrowth), terrain variables (slope, rocks, woody debris), operators' experience & skill and machinery limitations. The purpose of the study was to use remote sensing technology including LiDAR (Light Detection And Ranging) to identify and quantify tree size (m^3) and slope and hence derive relationships to predict productivity of a feller-buncher and a processor from remote sensing data.

A case study has been conducted in a 24 year-old radiata pine (*Pinus radiata*) plantation at Port Arthur, Tasmania, Australia. A time and motion study conducted in a clear-felling harvesting operation has been used to estimate the impact of tree size and slope on productivity of a feller-buncher and a processor. LiDAR flown in 2011 has been used to identify and quantify tree size and estimate slope. LiDAR-derived tree size estimates and slope have been ground-truthed against field measurements of trees and slopes respectively. Empirical models will be developed to enable the LiDAR-derived estimates of tree size and slope to be used to estimate productivity of harvesting equipment. The robustness of these relationships will be tested by applying the model to areas not used in the development process.

Key words: Harvesting system, Remote sensing, LiDAR, Productivity, Feller-buncher, Processor, Radiata pine.

Change detection of biomass along N-S 750 km transect in northwestern Canada by multi-temporal airborne laser profiling

Hayato Tsuzuki[1], Tatsuo Sweda[2] & Ross Nelson[3]

[1]Faculty of Agriculture, Ehime University, Japan zzukky@agr.ehime-u.ac.jp
[2]Professor Emeritus, Ehime University, Japan sweda2@hotmail.com
[3]Biospheric Sciences Branch, NASA-Goddard Space Flight Center
rfn104@gmail.com

Paper Number: SL2012-079

Abstract

Multi-temporal airborne laser missions over N-S transect of 750 km between Dawson (64.0 °N, 138.8 °W), Yukon territory and Inuvik (68.3 °N, 133.5 °W), northwest territory in Canada were conducted in summer of 2003 and 2011 for estimating terrestrial biomass change and monitoring how northern plants respond to possible global warming. By processing original surface profiles of the term-head and end profiling, vegetation profiles were obtained as a difference between surface and ground profiles. Then average vegetation height and timber volume/aboveground biomass every 100 m along transect at the term-head and end were obtained as average vegetation height and applying regression equations of timber volume/biomass measured at ground 80 sample plots against average vegetation height. Then change in timber stock and biomass in 8 years from 2003 to 2011 were calculated by the difference of those of term-head and end. Average vegetation height, timber stock and biomass entire 750 km transect were increased from 0.51 m, 15 m3/ha and 9.0 ton d.m./ha in summer of 2003 to 0.81 m, 23 m3/ha and 14.0 ton d.m./ha in summer 2011 as annual increment of 0.037 m, 1.1 m3/ha and 0.64 ton d.m./ha. Biomass increment was much larger in forest dominated section (0-70 km from Dawson; nearly 2 ton d.m./ha/year) than tundra dominated section (450-550 km from Dawson; 0.1 ton d.m./ha/year). Air temperature and warmth index in recent 25 years have been rising 0.34 degrees and 1.2 degrees-month (more than 0 degrees Celsius in mean monthly temperature) in Dawson, 0.63 and 0.3 in Eagle plains (the middle point of transect), and 0.78 and 1.9 in Inuvik respectively and which supports soil temperature should have also been rising in this area. Both the air and soil temperature rising causes the releasing growth limitation of plants in this region. However, biomass carbon release from forest fire seems not significant at least along our laser transect from year 2003 to 2011 judging from the combination of analysis of multi-temporal airborne laser profiling data set, fire history map from Yukon territory and observation from the air and ground. The cumulative vegetation height distribution curve which is one of the laser vegetation height indices obtained from height distribution of each 100 m section was used to clarify the mechanism of biomass change. The results revealed that the growth in northern plants within 8 years from 2003 to 2011 was basically by growing coverage by rather large trees (height more than 3 m) and growing height by stunted trees and grass (height less than 3 m) by sequestration of atmospheric carbon dioxide. These results may contribute to the basic information for understanding how northern plants have been and will correspond due to possible global warming.

Three-dimensional mass function clustering algorithm for crown segmentation

L. Monika Moskal[1], Nicholas R. Vaughn[2]

[1] School of Environmental and Forest Sciences, College of the Environment, University of Washington, Seattle, WA, USA
lmmoskal@uw.edu
[2] USDA Forest Service, Hilo, HI, USA nrv2@uw.edu

Paper Number: SL2012-080

LiDAR-based canopy height models are used for tree crown delineation derived with watershed algorithms, slope-based segmentations or even by using specialized segmentation software such as eCognition. These methods works well in plantations and forest with low heterogeneity and non-overlapping crowns, but have been shown to be problematic in heterogeneous forests. Moreover, as we begin working with full waveform LiDAR, it becomes obvious that the canopy height model is oversimplifying the canopy structure content of the data. Thus, approaches beyond surface structure are being developed, some working directly with the waveforms, or extracted point cloud data and others focusing on voxelizing the data. In this research a program was written that voxelizes the point cloud extracted from a full waveform dataset.
Our segmentation algorithm builds clusters of voxels by scanning through the voxel dataset one layer at a time. Within each layer each voxel containing points was either added to a neighbouring cluster if the cluster has enough "mass" over the voxel. Mass was computed as a combination of the sine of two angles: 1) between a vector connecting the voxel to cluster bottom and a vector connecting the voxel to cluster top, and 2) between the vertical plane and a vector from the voxel that was tangent to a circle of a radius computed from the current cluster width cantered at the cluster centroid. This algorithm was intended to allow crown outlines to overlap as they do in nature.

In order to assess accuracy, field measurements for 468 trees in a heterogeneous forest were utilized. To match the field data with the LiDAR-driven segmentation a spatial query was performed for each segmentation dataset matching all cluster outlines that contained a point representing the two-dimensional position of any tree. Clusters were allowed to match only one tree. Therefore in cases where multiple tree boles match a cluster, the tree with the smallest difference in crown width was selected. In the voxel-based segmentation, cluster outlines were allowed to overlap, making it possible for multiple clusters to match a single tree. In these cases, the cluster with the smallest difference in crown width was retained in the dataset. Then, an overall linear regression was performed, incorporating all matched trees and clusters of all species. The overall linear regression had an intercept of 1.41 and a slope of 0.81, with a t-value of -2.40. This yields a two-sided p-value of 0.017 on 313 degrees of freedom. The R2 statistic was 0.24. Furthermore, drastic differences in the R2 were observed for softwood versus hardwood species, indicating that the approach is sensitive to structural crown differences. Hardwood species are inherently difficult to segment, because of the lower apical dominance, also, multiple tops make a tree very easy to over-segment. Basic knowledge of the growth characteristics of species could improve the algorithm by incorporating species specific features such as average crown width into a segmentation algorithm. Such improvements serve as future direction for this research.
nick

Influence of stand and site conditions on the quality of digital elevation models underlying forests

Michael S. Watt[1], Thomas Adams[2], Pete Watt[3], Hamish Marshall[4], David Pont[5]

[1] Scion, Christchurch, New Zealand, michael.watt@scionresearch.com
[2] Metservice, Wellington 6012, New Zealand, thomas.adams@metservice.com
[3] Indufor Asia-Pacific Ltd, Auckland, New Zealand, pete.watt@indufor-ap.com
[4] Interpine, Rotorua, New Zealand, hamish.marshall@interpine.co.nz
[5] Scion, Rotorua, New Zealand, david.pont@scionresearch.com

Paper Number: ### SL2012-082

Abstract

When aerial LiDAR data is used to construct Digital Elevation Models (DEMs) under vegetation, DEM quality will invariably suffer due to attenuation of the laser pulses by the land cover. Although the ratio of ground returns to outgoing pulses (GR_{per}) is known to vary widely for forest applications, little research has quantified the influence of forest stand structure and site conditions on this ratio. An understanding of how these factors influence GR_{per} is crucial for the development of accurate digital elevation models. Using an extensive national dataset obtained from New Zealand's plantation forests the objective of this research was to develop a multiple regression model of GR_{per} that could be used to specify the necessary LiDAR point density for development of accurate DEMs.

Within the dataset GR_{per} averaged 30.5% ranging from 0.73 to 92.2%. The final model of GR_{per} included stand age, crop density, non-crop density and slope and accounted for 48% of the variance in GR_{per} with root mean square error (RMSE) of 13.9%. GR_{per} declined exponentially as stand age, crop and non-crop density increased and declined linearly with increases in slope. GR_{per} was not substantially affected by either pulse density, stand aspect or whether the stand comprised *Pinus radiata* or *Pseudotsuga mensiesii*.

The use of airborne and canopy lidar to validate continental-scale maps of vegetation height in Australia: A TERN project

Michael T. Schaefer[1, 3], Arancha Cabello-Leblic[1], Christopher Hopkinson[1], Alexander Held[1], Darius Culvenor[2], Matthew Paget[1]and Eva Van Gorsel[1]

[1]CSIRO Marine and Atmospheric Research, Canberra, GPO Box 3023, Canberra, ACT, 2601, Australia, Michael.Schaefer@csiro.au

[2]CSIRO Land and Water, Clayton, Victoria

[3]University of New England Precision Agriculture Research Group, University of New England, Armidale, NSW, 2351, Australia

Paper number: #SL2012-085

Abstract

The Australian Terrestrial Ecosystem Network (TERN) and its remote sensing data facility ("AusCover") and associated partners, aim to produce a national aboveground biomass map of Australia on an annual basis. Towards this end, a database of historical terrestrial laser scanning (TLS) data and airborne laser scanning (ALS) is being assembled, in particular to help in validating continental vegetation height estimates from across the Australia. The airborne lidar data will be aggregated to 25 m resolution, and compared, where appropriate, with the terrestrial TLS datasets, including those collected via the ECHIDNA validation instrument (EVI). Within a number of selected validation sites situated across Australia, additional *in situ* measurements of the vegetation height and stand characteristics will be used to provide more detailed structure information than is typically available within ALS and TLS datasets. These *in situ* measurements will enable the validation of the coinciding ALS and TLS datasets at each of the selected validation sites on a nationwide basis. Once validated, the ALS and TLS datasets themselves can be used to validate and calibrate national vegetation height products available via remote sensing such as ICESat GLAS data, which in turn will be used as inputs to the national biomass mapping project in TERN.

As per TERN data access principles, such datasets will be freely available, and would be valuable for future national and global scale vegetation canopy height validation and mapping activities, such as the launch of ICESat 2 in 2016, as well as for use in cross-validation activities of vegetation height and forest cover-change across Australia's ecosystems.

Contribution of ALS and TLS to structural profiles of complex, multi-strata cool temperate rainforest

Melissa Fedrigo[1], Kim Calders[2], Glenn Newnham[3], Darius Culvenor[3], Stephen Roxburgh[4], Sabine Kasel[1], Lauren Bennett[1] & Craig R. Nitschke[1]

[1]Department of Forest and Ecosystem Science; Melbourne School of Land and Environment; University of Melbourne; Burnley & Creswick, VIC; mfedrigo@student.unimelb.edu.au.
[2]Laboratory of Geo-Information Science and Remote Sensing; Wageningen University; Wageningen, The Netherlands; kim.calders@wur.nl.
[3]CSIRO Land and Water; Clayton South, VIC; Darius.Culvenor@csiro.au.
[4]CSIRO Sustainable Ecosystems; Canberra, ACT; Stephen.Roxburgh@csiro.au.

Paper Number: SL2012-089

Abstract

Airborne lidar (light detection and ranging) technology is increasingly being used for classifying and describing forests at landscape, regional and larger scales. Though sufficient in single-strata forests, the detail and accuracy of structural information from airborne lidar in complex forests with sub-canopy strata is limited by the overstorey gap fraction, the spatial density of lidar pulses, the number of return signals sampled per pulse, and occlusions.

Terrestrial lidar sensors allow for ground based scanning, which includes structural information of lower forest strata including coarse woody debris, but also extends into the higher strata. Increased structural information can be acquired by multiple terrestrial scan positions of the same plot to build detailed 3D reconstructions. However, terrestrial lidar is limited in complex forest structures by occlusion, and by reduced return signals when scanning dense upper canopy strata.

In this study, we aim to integrate airborne and terrestrial lidar data to provide more accurate representation of the vertical and horizontal structures of complex forests. Few studies have focused on the combination of both data sources to derive empirical relationships with structural attributes in complex forests. An analysis of the amount of information contributed by each lidar data source, and where in the forest profile each method contributes the most structural information, could be used to develop empirical relationships for predicting vertical structure and heterogeneity within complex forests.

We conduct a pilot study of the contribution of airborne and terrestrial lidar to structural profiles of multi-layered cool temperate rainforest in the Central Highlands of Victoria, south-east Australia. This forest type typically consists of a diverse rainforest tree layer beneath an emergent eucalypt canopy, includes multiple strata of ferns (including tree ferns) and has considerable coarse woody debris. In addition, the true extent of cool temperate rainforest in Victoria has not been accurately mapped, making it an ideal candidate for improved spatial characterisation. Airborne lidar data was acquired from November 2007 to January 2008 over the full extent of the Central Highlands. In addition, six cool temperate rainforest plots were scanned with a RIEGL VZ-400 3D terrestrial scanner in April 2012.

We provide a methodological framework for integrating airborne and terrestrial lidar to identify and characterise key structural elements in complex forest types like cool temperate rainforests. The structural variables demonstrate that potential empirical relationships can be derived from the combined lidar data to overcome the limitations of each individual method, and to improve forest structure classification at the landscape level.

Extraction of 3D tree models based on equirectangular projections of terrestrial laser scanning data

Lothar Eysn[1], Camillo Ressl[1], Andreas Grafl[1], Markus Hollaus[1], Werner Mücke[1]
Felix Morsdorf[2] & Norbert Pfeifer[1]

[1]Vienna University of Technology, Institute of Photogrammetry and Remote Sensing, Gusshausstraße 27-29, 1040 Vienna, Austria, le@ipf.tuwien.ac.at

[2]University of Zürich, Department of Geography, Remote Sensing Laboratories, Winterthurerstr. 190, CH-8057 Zürich, Switzerland, felix.morsdorf@geo.uzh.ch

Paper Number: SL2012-091

Abstract

Extracting 3D tree models based on high-density terrestrial laser scanning (TLS) point clouds with automatic, semi-automatic or manual methods is a challenging task as trees are complex, individual objects. Various publications in this research field show the demand for tree reconstruction methods based on TLS data. Data sets acquired with current TLS devices allow detailed reconstruction of tree stems and branches, which are a fundamental input for e.g. stem volume assessment or setting up virtual forest scenes. For dense forests, a completely automated reconstruction of a tree is often limited by occlusions and data gaps, as well as varying point density. Beside automatic extraction methods based purely on the point cloud (e.g. region growing algorithms), tree models can be semi-automatically created by e.g. local cylinder fitting. This task is challenging and time consuming because the interpreter has to navigate through dense point clouds and the selection of subsets for cylinder fitting can be tricky.

In the presented approach a semi-automatic method for extracting coniferous and deciduous tree models based on projected 2D maps of the TLS point cloud is described. Equirectangular projections (EP) based on the observation angles of the scan are created, thereby displaying the distance (range map, RM) and intensity information (intensity map, IM) detected by the scanning device. The so-called tree structure elements (i.e. stems and branches) are clearly interpretable in the IM and RM. These easily navigable maps provide a good basis for extracting trees by digitizing the axis of the structure elements and assigning their respective local diameter which were measured in the 2D maps. EP derived from multiple scan positions around the trees are used to complete occluded sections. Erroneous measurements, arising from moving tree parts (e.g. branches affected by wind), or by imperfections in the relative orientations of the scans, are overcome because the extraction of the tree structure is performed using single maps instead of a merged point cloud of individual scans. The digitized 2D skeletons are transformed to 3D space and furthermore extruded to 3D models. Additionally to the modelling process a classification of the point cloud into two classes (stem/branches, needles/leaves) is performed using a voxel approach. The tree models are a fundamental input for this classification.

The method is applied to a dense TLS dataset acquired in a managed forest in Tharandt, Germany. About 34 scans were carried out during the data acquisition to measure approximately 90 spruce and fir trees with minimal occlusions. The results demonstrate the feasibility of extracting tree models semi-automatically based on 2D maps with a very high degree of completeness. In comparison to other approaches, the number of reconstructed trees is higher (by factor 3) than the number of scans. The quality assessment was based on a comparison of the point cloud and the cylinder model's 3D views and quantitative evaluation. Visual assessment showed deviations in the order of the measurement accuracy and tree surface irregularity. The extracted tree models are used to set up a virtual forest scene for radiative transfer modelling.

Relations between Airborne Laser Scanning-metrics and Lorenz curve descriptors for forest structure characterization

Rubén Valbuena[1], Matti Maltamo[2], José Antonio Manzanera[3], Petteri Packalén[2], Francisco Mauro[3], Lauri Mehtatalo[2], Cristina Pascual[3]& Gert-Jan Nabuurs[1]

[1]European Forest Institute HQ. Torikatu 34. 80100 Joensuu, Finland; ruben.valbuena@efi.int
[2]University of Eastern Finland. Faculty of Forest Sciences. PO Box 111 Joensuu, Finland; matti.maltamo@uef.fi; petteri.packalen@uef.fi; lauri.mehtatalo@uef.fi
[3]Technical University of Madrid. School of Forestry. Ciudad Universitaria s/n. 28040 Madrid, Spain; joseantonio.manzanera@upm.es; francisco.mauro@upm.es; c.pascual@upm.es

Paper Number: SL2012-094

Abstract

Lorenz ordering has been found to be a reliable method for discriminating even from uneven-sized forest areas. For this reason, we considered using Gini Coefficient (GC) and Lorenz asymmetry (S) with the intention of providing a bivariate description of forest structure. We subsequently considered that a prediction of this response by means of Airborne Laser Scanning (ALS) can provide wall-to-wall characterization of forest structure. The first step was consequently to study the relations between ALS-metrics with this response. Exploratory multivariate analysis was carried out by following a information-theoretic approach for multimodel inference (MMI) and partial least squares (PLS). The explained variance was studied by means of the relative variable importance obtained in the MMI averaged model. The different components of predictor/response relations were evaluated from the weights of diverse significant PLS latent variables, as well as the overall variable importance parameter.

Results showed the importance of considering Lorenz asymmetry from the point of view of the combined effect of differences in both stem density and basal area relative to quadratic mean diameter (Dg). With regards to the relation between ALS-metrics and the GC of inequality among tree diameters, it was demonstrated that variables explaining most variance were canopy coverage (Cov; percentage of returns above one metre), standard deviation (SD), also with high levels of relative importance for the maximum (Max) and the tenth percentile (P10). The relative amount of basal area stocked above Dg was mainly related with the L-moment ratio of kurtosis (L-kurt) and SD, i.e. the shape and dispersion of the ALS heights histogram. On the other hand, using combinations of lower and higher percentiles (P10 and P90) was the best way to estimate the proportion of stem density accounted by the trees of diameter bigger than Dg, with a high relative importance also for the concentration of ALS-heights (L2).

These results suggest that ALS surveys can be a reliable method for estimating the inequality among tree sizes. GC discriminated the uneven-sized groups with high separability from the even-sized plots, therefore having small probability for confusion. S can be used as a means of differentiating areas showing a reverse J-shaped diameter histogram against those having a bimodal distribution. The overall method seems reliable for studying the irregular spatial distribution of several forest structural types. One drawback found was the dependency of the response on the scale used, as the method has an implicit bias towards considering a plot more uneven-sized when increasing its size. Further research should therefore be focused on effects of plot size and the spatial resolution considered for area-based methods in ALS forest inventory.

Investigations on the orientation of terrestrial laser scanning point clouds in dense forests

Lothar Eysn[1], Werner Mücke[1], Camillo Ressl[1], Markus Hollaus[1], Reik Leiterer[2], Franz Blauensteiner[1] & Norbert Pfeifer[1]

[1] Vienna University of Technology, Institute of Photogrammetry and Remote Sensing, Gusshausstraße 27-29, 1040 Vienna, Austria, [le, wm, car, mh, np, fb]@ipf.tuwien.ac.at

[2]University of Zürich, Department of Geography, Remote Sensing Laboratories, Winterthurerstr. 190, CH-8057 Zürich, Switzerland, reik.leiterer@geo.uzh.ch

Paper Number: SL2012-097

Abstract

Vegetation structure was identified to have significant influence on the course of environmental processes. For complex applications in environmental context, such as radiative transfer modelling in forests, highly detailed and accurate geometric depiction of the vegetation is among the basic requirements to ensure as realistic conditions as possible for the modelling algorithms. Terrestrial laser scanning (TLS) provides reliable means for collecting the required 3D information for this purpose. Especially at ranges not in the immediate vicinity of the viewpoints, the direct 3D collection is advantageous to the indirect acquisition via photographs, due to the many occlusions in the forest environment. Study sites in dense mountainous forests pose a special challenge because of the complexity of the terrain and the vegetation. Detailed a-priori information is typically not available, and therefore the design of the measurement campaign often has to be arranged on the spot. Undersampling of the low vegetation and minimal overlap between scans for reduced time consumption in the field, as well as wind limit the use of ICP as orientation method of choice. Positioning and setup of reference targets, which are therefore required for merging the data from different scanning positions, can become a tedious and time-consuming task.

Within this study we compare two different approaches for the orientation of point clouds from individual scans acquired with a phase-shift scanner in two densely forested areas in Tharandt, Germany and Lägern, Switzerland. The study area in Switzerland consists of a steep sloped, unmanaged, forest with various tree species. The study area in Germany consists of a managed forest with very old spruce and fir trees. A network of reference points was set up and measured with a total station for each study area to serve as basis for the investigations. These reference networks consisted of commercially available planar precision targets and self-designed spherical targets made from lightweight materials. These spherical targets have considerable advantages over conventional ones. Due to their size and reflective properties the targets feature increased visibility under leaf-on conditions, their shape offers high flexibility for the positioning of the laser scanner and they are designed to stand alone. They can be located anywhere without the need to mount them on a near-by tree. The planar targets, which display individual black-and-white patterns, were mounted on trees. The first approach uses all available spherical targets for orientating the individual scans. In contrast, for the second method the laser scanner was equipped with two distinct standard reflector prisms for measurement with a total station. This setup allows direct georeferencing and orientation of each individually acquired TLS scan without the need for reference targets.

Both methods are compared on the basis of practicability of data acquisition, as well as relative and absolute geometric accuracy of the resulting point clouds. The set of mounted planar precision targets is used for evaluation purposes.

Detection of young regenerations in temperate forest areas attacked by bark beetle by exploiting airborne full-waveform LiDAR data

Wei Yao[1], Peter Krzystek[1] & Marco Heurich[2]

[1]Department of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany, {yao, Krzystek}@hm.edu
[2]Bavarian Forest National Park, 94481 Grafenau, Germany, marco.heurich@npv-bw.bayern.de

Paper Number: SL2012-098

Abstract

LiDAR remote sensing has been developing as a promising additional instrument to the conventional field work towards forest inventory. Unfortunately, understory trees such as small regenerations in multilayered forest stands are often ignored. However, Information about post-disease or natural forest regeneration and will provide insight on dynamics of future stand composition and structure. Information about them would also be valuable to wood procurement and biodiversity management. Therefore, a detailed understanding of the availability and spatial distribution of forest regenerations is required.

Currently, field-based inventory is the principal means used to collect information regarding regeneration status, since it is very time-consuming and based on plot sampling which is only sparsely distributed. The full-waveform LiDAR with high penetration ability against overstory crowns can give us new hope to resolve regenerations. In this work, we want to examine the potential of the small-footprint full-waveform laser scanner for the mapping of forest regeneration in temperate mountain forests which are recovered from an attack from bark beetles. Usually, the overstory prevents from obtaining a wall-to-wall sample of understory, and laser measurements are subject to transmission losses. Therefore, discriminating between the crowns of dominant and suppressed trees in LiDAR point clouds is very challenging for both the sensor technology and methodology.

This paper is to present an automatic strategy for inferring the spatial distribution of regenerations in a structurally inhomogeneous forest area by extracting 3D individual trees from airborne LiDAR data. Regeneration trees are obtained based on classifying the single trees with object-level features which are detected by 3D normalized cut segmentation. The referenced data for regenerations were manually either cropped out from or marked in very high-density TLS data of the same sites, which are deemed to be adequate for serving as ground truth. We applied the strategy to two sample plots acquired under leaf-on and leaf-off conditions and validate the results by two different methods: single-tree based and area-based evaluations. The single tree evaluation is performed based on single-tree level by finding matched tree segments in the reference, while the area-based evaluation is performed by area (pixel)-wise (1 meter grid cell) comparing the detected regenerations to the ground truth in the horizontal (x-y) dimensions.

We assessed tree regeneration in central-Europe mixed spruce stands and experienced more than 50% overstory mortality due to bark beetles. We presented results that show differences in performance between two validation methods that use different reference data and comparison mechanisms. For both test sites, we have obtained a detection rate of up to approx.45% with 25% false alarms based on single-tree evaluation. However, when using the area-wide evaluation we have achieved an improved result, namely up to nearly 60% for the detection rate with 18% false alarms.

To the best of our knowledge, this study made the initial attempt to characterize the distribution of 3D individual regenerations in heterogeneous forests from airborne LiDAR and has demonstrated for the first time the feasibility of single tree detection from remotely sensed data for forest regeneration analysis.

References

Chapelle, O., Vapnik,V., Bousquet, O. and Mukhherjee, S., 2002. Choosing multiple parameters for support vector machine. *Machine Learning*, 46 (1-3), 131–159.

Collins, B J., Rhoades, C C., Hubbard, R M. and Battaglia, M A., 2011. Tree regeneration and future stand development after bark beetle infestation and harvesting in Colorado lodgepole pine stands. *Forest Ecology and Management*, 261 2168-2175.

D'Errico, 2006 D'Errico, J., 2006, Surface fitting using gridfit. http://www.mathworks.com/matlabcentral/fileexchange(accessd on 10.01.2009)

King, D J., Pitt, D G. and Pouliot, D A., 2005. Development and evaluation of an automated tree detection–delineation algorithm for monitoring regenerating coniferous forests. *Canadian Journal of Forest Research*, 35(10), 2232-2345.

Korpela, I., Ørka, H.O., Maltamo, M., Tokola T. and Hyyppä, J., 2010. Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *Silva Fennica*, 44(2), 319–339.

Korpela, I., Hovi, A. and Morsdorf, F., 2011. Mapping understory trees using airborne discrete-return LiDAR data. *ISPRS Hannover Workshop*, June , 2011, on CD-rom.

Li, A., Huang, C., Sun, G., Shi, H, Toney, C., Zhu, Z., Rollins, M., Goward, S. and Masek J., 2011. Modeling the height of young forests regenerating from recent disturbances in Mississippi using Landsat and ICESat data. *Remote Sensing of Environment*, 115, 1837-1849.

Næsset E., 2011. Estimating above-ground biomass in young forests with airborne laser scanning. *International Journal of Remote Sensing*, 32(2), 473-501.

Potter, C., Li, S., Huang, S. and Crabtree, R L., 2012. Analysis of sapling density regeneration in Yellowstone National Park with hyperspectral remote sensing data, *Remote Sensing of Environment*, 121, 61-68.

Reitberger, J., Krzystek, P. and Stilla, U., (2008). Analysis of full waveform LiDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing*, 29(5), 1407 – 1431.

Reitberger, J., Schnörr, C., Krzystek, P. and Stilla, U., 2009. 3D segmentation of single trees exploiting full waveform LiDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), 561–574.

Shi, J. and Malik, J., 2000. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888 – 905.

Stilla, U., Yao, W. and Jutzi, B., 2007. Detection of weak laser pulses by full waveform stacking. In: Stilla U, et al (eds) PIA07 Photogrammetric Image Analysis 2007. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, Vol 36(3/W49A):25-30.

Yao, W. and Stilla, U., 2010. Mutual Enhancement of Weak Laser Pulses for Point Cloud Enrichment Based on Full-Waveform Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 48(9), 3571-3579.

Yao, W., Krzystek, P. and Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. *Remote Sensing of Environment*, Vol.123, 368-380.

Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M. and Viitala, R., 2011. Predicting individual

tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(1), 28-37.

Acknowledgement

The work is also supported by Open Fund of State Key Laboratory of Remote Sensing Science (Grant No. OFSLRSS201212)

Tree height distribution estimation from ALS data in Central Spain.

Francisco Mauro[1], Esperanza Ayuga Téllez[1] Concepción González[1] & María Victoria Nuñez[1] Maria Angeles Grande[1]

[1] ETSI Montes. Polytechnic University of Madrid. Ciudad Universitaria SN. 28040 Madrid francisco.mauro@upm.es

Paper Number: SL2012-100

Height estimates are among the most accurate parameters generated from ALS data. However, most of the studies estimate aggregated values such as mean height or dominant height. A few experiences have explored the possibilities of estimating the tree height distributions (THD) which provide information about size variability and forest structure. These experiences relay on maximizing the likelihood of ALS point clouds and/or require a previous knowledge about crown shapes. Although these approaches are very robust from a theoretical perspective, they might be too complicated for standard inventories. Different, researchers have estimated diameter at breast height distributions using: the well known Area Based Approach (ABA) and Parameter Prediction Methods or Parameter Recover Methods. These approaches don't have a deep theoretical background about the interaction of ALS pulses and the forest canopy. However, they are simpler and provided good results. To the date no similar study has been applied for estimating THD. The aim of this study is analyzing the performance of two different ABA methods for estimating THDs. The first method M1 directly estimates the number of stems in five height intervals. The second one M2 recovers THD parameters from estimates of stand density N and mean and Dominant Height (H^{-},H_0).

The study area is a 300-ha Pinus sylvestris forest located in Central Spain. Tree Height was measured for every tree in a total of 37 GPS georeferenced circular plots of 20 m radius. For the M1 estimation method several combinations of tree height interval classes were defined and the number of trees within each interval was computed for every plot. Multiple regression models relating ALS data and number of stems were fitted for each height interval. The selected intervals ((1.30,4], (4,10], (10,16], (16,25], (25,∞) units in meters) were those with the highest R2 value that also allowed a complete cover of the range of heights. Plot level estimated values of the number of stems within these intervals were compared to field values. For the M2 method, linear regression models were fitted for predicting N, H^{-} and H_0 form ALS predictors. A two parameter Weibull model was chosen to describe THDs. A system of equations can be established for obtaining α,β from N, H^{-},H_0. Plot level estimates of these variables provided by their predictive models were used for estimating plot level values of α and β. The obtained THDs were scaled to the estimated N and discretized in the mentioned intervals. These THDs were also compared to Maximum Likelihood ML estimates.

The M1 method provided the best estimates being the RMSE for each interval 120.8, 107.4, 83.9, 102.1, 37.4 stems per hectare. The R2 were 85.2%, 84.4%, 83.4%, 86.93%, 91.6%. M2 accounted for less than 20% of the variance for the second and third height class. No variance was explained for the other intervals. Significant differences between parameters estimated using M2 or ML were found. The first method seems to be a good way for estimating THDs. The second method should be tested after stratification as it didn't provide realistic estimates when no stratification was performed.

3D-mapping of tree spectral indices using hyperspectral LiDAR

Teemu Hakala[1], Juha Suomalainen[1], Olli Nevalainen[1], Eetu Puttonen[1] & Sanna Kaasalainen[1]

[1] Department of remote sensing and photogrammerty, Finnish Geodetic Institute, P.O. box 15, 02431 Masala, Finland, Contact: teemu.hakala@fgi.fi

Paper Number: #SL2012-102

Abstract

Recent advances in nonlinear fiber optics have led to the development and commercial availability of extremely broadband laser sources. The number of applications has increased in the recent years in, e.g., biomedical optics or chemical analysis. Combined with a hyperspectral time-of-flight sensor, the supercontinuum laser sources can be used for simultaneous measurement of distance and reflectance spectrum, which has been the basis for a prototype full waveform active hyperspectral LiDAR developed at the Finnish Geodetic Institute. We have investigated the potential of this novel instrument for remote sensing applications, such as mapping tree spectral indices from hyperspectral 3D point clouds. Our prototype can measure both the range and the spectral information from a single laser pulse. The spectra can be used in visualization and automated classification of the laser scanner point cloud and the retrieval of spectral indices and target physical properties. At this stage the instrument optimized for short-range terrestrial applications.

In this paper, we present an application of this novel technique into the measurement of 3D distribution of chlorophyll index within a tree canopy. We have calculated modified chlorophyll absorption ratio index (MCARI) from the spectral point cloud to give us an indication about how the chlorophyll is distributed within the tree canopy. Chlorophyll gives an indication about the photosynthesis activity and the nutrition state of the tree, thus allowing healthy branches to be detected from branches in poor state. Additionally, multitude of other spectral indices can similarly calculated from the dataset to show e.g. how the moisture content varies inside the target.

We have tested this method on two species of coniferous trees; scots pine (Pinus sylvestris) and Norway spruce (Picea abies) with 10 specimen of each measured inside laboratory. Samples of each tree were taken and the chlorophyll content of these samples was measured. Validation was carried out by means of traditional laboratory measurements of chlorophyll content.

Science and technology towards precision forestry

Markus Holopainen[1*], Juha Hyyppä[2], Mikko Vastaranta[1] & Hannu Hyyppä[3]

[1]University of Helsinki, Department of Forest Sciences, Finland –first.last@helsinki.fi
[2]Finnish Geodetic Institute –first.last@fgi.fi
[3]Aalto University, Research Institute of Modelling and Measuring for the Built Environment, Finland –first.last@aalto.fi

Paper Number: SL2012-109

Abstract

In 2008-2011 our research group had a Academy of Finland project "*Improving forest supply chain by means of advanced laser measurements" (L-impact*) and in 2011-2014 we have an on-going project "Science and technology towards precision forestry" (PreciseFor). During these projects new methodologies have developed in the fields of laser scanning technology and forestry applications. In L-impact and PreciseFor projects, our group has published over 70 scientific papers. Our research has focused on novel technologies, methods and algorithms in airborne -, terrestrial - and mobile laser scanning (ALS, TLS, MLS) including UAV. We have developed beyond state-of-art applications for forest inventory, monitoring, mapping, management, and planning. Our scale has varied from tree-level to large-area applications. In this poster we summarize some highlights of the L-impact and PreciseFor projects during 2008-2012 including uncertainties in forest management planning calculations, novel individual tree detection methods, accuracy of TLS and MLS in tree mapping. and fusion of 3D information derived from ALS and SAR radargrammetry (e.g. Holopainen et al. 2010a,b,c, Hyyppä et al. 2012, Karjalainen et al. 2012, Liang et al. 2012, Lin et al. 2012, Vastaranta et al. 2011, 2012a, 2012b, Yu et al. 2011).

References

Holopainen, M., Mäkinen, A., Rasinmäki, J., Hyyppä, J., Hyyppä, H., Kaartinen, H., Viitala R., Vastaranta, M. & Kangas, A. 2010a. Effect of tree level airborne laser scanning accuracy on the timing and expected value of harvest decisions. European Journal of Forest Research, (2010) 129:899-910.

Holopainen, M., Mäkinen, A., Rasinmäki, J., Hyytiäinen, K., Bayazidi, S. & Pietilä, I. 2010b. Comparison of various sources of uncertainty in stand-level net present value estimates. Forest Policy and Economics, 12(2010):377-386.

Holopainen, M., Haapanen, R., Karjalainen, M., Vastaranta, M., Hyyppä, J. Yu, X., Tuominen, S. & Hyyppä, H. 2010c. Comparing accuracy of airborne laser scanning and TerraSAR-X radar images in the estimation of plot-level forest variables. Remote Sensing. 2010, 2:432-445.

Hyyppä, J., Yu, X., Hyyppä, H., Vastaranta, M., Holopainen, M., Kukko, A., Kaartinen, H., Jaakkola, A., Vaaja, M., Koskinen, J. & Alho, P. 2012. Advances in Forest Inventory Using Airborne Laser Scanning. Remote Sensing 2012, 4, 1190-1207.

Karjalainen, M., Kankare, V., Vastaranta, M., Holopainen, M. & Hyyppä, J. 2012. Prediction of plot-level forest variables using TerraSAR-X stereo SAR data. Remote Sensing of Environment 117: 338—347.

Liang, X, Litkey, P., Hyyppä, J., Kaartinen, H., Vastaranta, M. & Holopainen, M. 2012. Automatic stem-mapping using single-scan terrestrial laser scanning. IEEE Transactions on Geoscience and Remote Sensing, 50:661-670.

Lin, Y, Hyyppä, J., Jaakkola, A. & Holopainen, M., 2012.Characterization of mobile LiDAR data collected with multiple echoes per pulse from crowns during foliation. Scandinavian Journal of Forest Research, 27:3, 298-311.

Vastaranta, M., Holopainen, M., Yu, X., Hyyppä, J., Hyyppä, H. and Viitala, R. 2011. Predicting stand-thinning maturity from airborne laser scanning data, Scandinavian Journal of Forest Research, 2011(26):187-196.

Vastaranta, M., Kankare, V., Holopainen, M., Yu, X., Hyyppä, J. & Hyyppä, H. 2012a. Combination of individual tree detection and area-based approach in imputation of forest variables using airborne laser data. ISPRS Journal of Photogrammetry and Remote Sensing 67: 73—79.

Vastaranta, M., Korpela, I., Uotila, A., Hovi, A. & Holopainen, M. 2012b. Mapping of snow-damaged trees in bi-temporal airborne LiDAR data. European Journal of Forest Research.

Yu, X., Hyyppä, J., Vastaranta, M. Holopainen, M. & Viitala, R. 2011. Predicting individual tree attributes from airborne laser point clouds based on random forest technique. ISPRS Journal of Photogrammetry and Remote Sensing 66:28-37.

Improving the vegetation classification accuracy with height information from ICESat/GLAS

Huabing Huang[1], Peng Gong[1] , Caixia Liu[1], Xiaoyi Wang[1], Congcong Li[1], Peng Zhu[1]

[1] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Beijing, 100101, China; huanghuabing@irsa.ac.cn

Paper Number: SL2012-112

The ICESat/GLAS was used to assess forest aboveground biomass to a large extent as the unique remote sensing data, which can provide the vegetation height information globally. It is hard to classify forest and shrub with high accuracy when using optical image, solely. So the vegetation height information should be considered when mapping global land cover and land use with an automatic method. In this paper, we choose five different areas at the continental scale, and calculate the vegetation height of forest, shrub and grass. The separability of three vegetation types using vegetation height was evaluated. The samples were choose from changbai mountain in northeast China, Sierra Nevada forest in USA, Rondonia forest in Brazil, boreal forest in Canadian and tropical dry forest in Africa and 3,017 samples were choose by visual interpolation with Google Earth. The height was determined from ICESat/GLAS terrestrial product named GLA14. The statistical results show that there is an obvious separability between forest, shrub and grass. The shrub is mixed with grass in some extent. So more work and height extracting method should be done to acquire better vegetation height in order to classify vegetation using height information.

References:

Baccini,A., et al., (2011), Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Climate Change, 2, 182–185

Chen, Q. (2010). Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific Coast region using satellite laser altimetry. Remote Sensing of Environment, 114, 1610–1627

Dolan, K., J. G. Masek, C. Huang, and G. Sun (2009), Regional forest growth rates measured by combining ICESat GLAS and Landsat data. J. Geophys. Res., 114, G00E05, doi:10.1029/2008JG000893.

G. Sun and K.J. Ranson (2000), Modeling lidar returns from forest canopies, IEEE Transactions on Geoscience and Remote Sensing, 38(6), pp.2617- 2626.

Lefsky, M. A. (2010), A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System, Geophys. Res. Lett., 37, L15401.

Saatchi, S.S., et al., (2011), Benchmark map of forest carbon stocks in tropical regions

across three continents. Proceedings of the National Academy of Sciences, 1-6.

Simard, M., N. Pinto, J. B. Fisher, and A. Baccini (2011), Mapping forest canopy height globally with spaceborne lidar, J. Geophys. Res., 116, G04021.

Estimation of forest parameters in Northern subtropical secondary forest, China, using airborne LiDAR

Cao L.[1], Coops N.C[2], She G.[3], Dai J.[4]

[1]University of British Columbia ginkgocao@gmail.com
[2]University of British Columbia nicholas.coops@ubc.ca
[3]Nanjing Forestry University ghshe@njfu.edu.cn
[4]Nanjing Forestry University Aaron-dai@163.com

Paper Number: #SL2012-113

Abstracts

Light Detection and Ranging (LiDAR) is becoming an important tool in forestry around the world due to its ability to accurately estimate forest structure variables and biophysical properties. In this study, we demonstrate the application of airborne LiDAR technology, using regression models, to estimate forest structural metrics and biomass components for a forest farm in Jiangsu province, China.

In June - July, 2012, 45, 30 x 30m plots in Yushan forest farm, Jiangsu province, China were established, which is a typical Northern sub-tropical hilly secondary forest in the Middle-lower Yangtze plain. The plots are located in homogeneous forest stands of Masson pine(*Pinus massoniana*), Chinese fir(*Cunninghamia lanceolata*), Sweetgum(*Liquidambar formosana*), German oak(*Quercus acutissima*), and so on, resulting in three broad vegetation types: coniferous (10 plots), broadleaf (13 plots) and mixed (22 plots). At each plot DBH and tree height were measured and plot-level summaries such as basal area, Lorey's mean height, as well as biomass components (including stem, branch and foliage biomass) calculated using local algometric equations.

A suite of statistical models based on LiDAR-derived percentiles were developed to predict the biomass components; and tests undertaken to assess the accuracy and improvement of LiDAR-derived biomass model compared with ground-survey based empirical biomass model.

Unsupervised training for treetop detection with airborne laser scanning data

Jean-Matthieu Monnet[1], Frédéric Berger[1] & Jocelyn Chanussot[2]

[1] Irstea, UR EMGR, F-38402 St-Martin-d'Hères, France (jean-matthieu.monnet, frederic.berger@irstea.fr)
[2] GIPSA-Lab, Grenoble Institute of Technology, BP 46, 28402 Saint Martin D'Heres, France (jocelyn.chanussot@gipsa-lab.grenoble-inp.fr)

Paper Number: SL2012-114

Abstract

Numerous methods have been proposed for the detection of single trees in airborne laser scanning (ALS) data. Most of them are highly dependent on the initial settings of the algorithm, as parameters (e.g. smoothing of the digital canopy height model) will affect the overall detection performance, and more particularly the trade-off between omission and commission errors. To tackle this issue, the use of prior information about the forest stand is possible when ground truth for the area is available. Alternatively, adaptive parametrization in the course of the detection procedure requires more complex algorithms which might have trouble when processing large areas.

In this article a procedure for automated parametrization is presented. It is based on the unsupervised training of the detection algorithm with reference forest plots including coregistered field and ALS data. The local maxima filtering algorithm is adopted as it is simple and fast. The training step consists in evaluating the detection performance of the algorithm on the reference plots for several parameter combinations. Detection quality is evaluated as a trade-off between the number of correctly detected trees and the number of false detections. When trees are to be detected in a newly surveyed area, two possibilities for algorithm parametrization are compared. The first option is to use the parameter combination that is the more robust when used on the training set ("average" setting). The second option is to use the combination that yields the best detection on the ALS point cloud from the training set that best resembles the new data. The matching criterion is based on the Fourier spectrum of the canopy height model computed from the point cloud.

26 forest plots located in seven different ALS surveys of mountainous areas are used to test the workflow. Plots have a minimum area of 0.25 ha and represent various stand structures and tree species. To compare the two parametrization options and evaluate their sensitivity to the training set size (number of reference plots), a cross validation procedure based on repetitive sampling of the training set among the available 26 plots is performed. Results show that for training sets with less than 15 plots, the "average" setting performs better, whereas with training sets larger than 20, the matching procedure yields better detection performance.

This method for unsupervised training is quite flexible as it can be used with any detection algorithm that requires initial parametrization. Moreover, the detection performance criterion can be modified in order to reflect the end-user preference regarding detection results. This study is an example of how single tree methods can benefit from an area-based analysis. Further work should investigate whether metrics usually computed for area-based methods (e.g. height quantiles) could also improve the point cloud matching.

Estimation of crown volume from airborne lidar data using alpha shapes

Lauri Korhonen[1], Jari Vauhkonen[2], Anni Virolainen[3], Aarne Hovi[4] & Ilkka Korpela[5]

[1]University of Eastern Finland, School of Forest Sciences lauri.korhonen@uef.fi
[2]University of Helsinki, Department of Forest Sciences jari.vauhkonen@helsinki.fi
[3]University of Eastern Finland, School of Forest Sciences anni.virolainen@uef.fi
[4]University of Helsinki, Department of Forest Sciences aarne.hovi@helsinki.fi
[5]University of Helsinki, Department of Forest Sciences ilkka.korpela@helsinki.fi

Paper Number: SL2012-115

Abstract

Crown volume (CV) is needed in many theoretical forest models. The problem is that direct measurement of CV and crown shape is difficult using conventional field methods, and therefore allometric relationships and simple geometrical shapes are commonly used. However, airborne lidars and methods of computational geometry enable direct calculation of individual crown properties from the 3D triangulation of the lidar point cloud. Our objective was to compare the CVs obtained this way with field-measured values. The field data included 35 Scots pines, 25 Norway spruces, and 17 silver birches in Hyytiälä, Finland. The control CVs were obtained by measuring the crown radii at different heights using a specific instrument, the angle measurer. It is a T-shaped stick, which has an angular scale with one degree resolution. Each tree was measured from two perpendicular directions, so four profiles were available per tree. The CV was calculated by integrating the individual line segments as solids of revolution and averaging the derived volumes.

A digital surface model (DSM) describing canopy height was created from the lidar data. Then the trees were detected from the DSM using watershed segmentation. The segments were used as a basis for intersecting the echoes belonging to the crown, but the final echo selection was done manually as some of the segments had unrealistic shapes or contained several trees. An automated method for detecting the height of the crown base (CBH) was applied to omit echoes from the ground and the understory. Alternatively, the field-measured CBH was used as prior information to eliminate this error source. Finally the CV was extracted from the triangulation of the point data using 3D alpha shape and 3D convex hull techniques.

The lidar-derived CVs were generally smaller than the field-based estimates. The best correspondence with the field-measured CV was obtained using the convex hull of the point data and field-measured CBH (RMSE 42.7%; bias 21.7%). In an attempt to describe the crowns in more detail we also used pseudo-optimal alpha value as a size-criterion for the cells of the triangulation, but this way the CV was underestimated more considerably (RMSE 80.4%; bias 56.1%). The CBH was usually overestimated from the lidar (RMSE 39.7%; bias -22.4%), especially for spruces that had long crowns. Thus using the lidar-derived CBH also increased the bias of the CV estimates (convex hull RMSE 60.1%; pseudo-optimal alpha RMSE 88.7%). Nevertheless, the results indicate that if the echoes belonging to the crown can be delineated correctly, the crown volume can be directly estimated from the lidar point cloud. One reason for the frequent underestimation was that the angle measurer recorded the extreme edges of the crown envelope, while lidar echoes usually penetrate slightly into the crown and do not reliably map the extreme dimensions. Also the number of echoes was often inadequate in the lower parts of the crowns.

Invention of a method to use airborne LiDAR intensity and DCHM data for forest type classification

Katsumasa OONO[1], Fumihiko ITO[2] & Satoshi TSUYUKI[1]

[1]Graduate School of Agricultural and Life Sciences, The University of Tokyo
katsumasa2501@gmail.com
[2] Asia Air Survey Co.,Ltd. fmh.itoh@ajiko.co.jp

Paper Number: SL2012-116

Abstract

Aerial photograph is a data source widely used for vegetation analysis. There are two methods for image analysis – manual interpretation by human eyes and automatic classification by using software. From the point of view of work efficiency, automatic classification is preferable. However, aerial photograph is not so suitable for automatic classification of vegetation. The reason is because even if vegetation type is the same, its tone appeared on photographic image is not homogeneous if image capture date, slope direction or slope angle is different. This problem occurs because aerial photography captures energy of sun light reflected by earth surface objects. Recently, LiDAR is widely used in the measurement of earth surface undulations in detail. Data recorded by LiDAR sensor are not affected by the strength of reflected sun lights. LiDAR uses laser beam to scan the earth surface and its beam is in the spectrum range of near infrared with 1,064nm wave length. Because of this wave length, there is a possibility that intensity of reflected laser pulse can be used to classify vegetation types. In this paper, we devise Laser Forest Type Image which is used as an alternative to aerial photograph and clarify the characteristics of Laser Forest Type Image. Our study area is the forested area of Saga Prefecture, Japan, which covers approximately 1,800 km^2. LiDAR data acquisition was conducted in the summer of 2011 at 4 points/m^2 point density. Laser Forest Type Image is generated by combining three kinds of LiDAR data. They are, (1) intensity of laser pulse, (2) DCHM (Digital Canopy Height Model) calculated from difference between DCSM (Digital Canopy Surface Model) and DTM (Digital terrain Model) and (3) canopy shape derived from DCHM. According to published research papers, laser pulse intensity of pine and spruce is different. In our own field survey, it was found that the difference in the intensity of laser pulse is related to the difference in the size of leave area such as LAI. Then, Laser Forest Type Image was colored so that three types of species – Japanese cedar, Japanese cypress and others can be identified. Laser Forest Type Image is suitable for automatic classification because tone of land cover is homogenous regardless of date of data capture, slope angle or slope direction if land cover type is the same. Further, in case of Laser Forest Type Image, vegetation is displayed in wider spectrum of color compared to aerial photo in which vegetation is shown basically in green colors only. As a test of automatic classification, land cover classification was done by an object based method using e-Cognition. Targets of the classification were Japanese cedar, Japanese cypress, broadleaves and others. From the verification of the classification results, the overall accuracy is higher than 85%. This result is considered that vegetation maps can be produced more efficiently compared to visual interpretation of aerial photos. Laser Forest Type Image could be utilized as a base data for vegetation mappings.

Reference

Daniel N.M. Donoghue, Peter J. Watt, Nicholas J. Cox and Jimmy Wilson, 2007. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment*, 110, 509-522.

Using LiDAR to predict hiding cover – an important determinant of predation risk of roe deer

Karen Lone, Terje Gobakken, Atle Mysterud, Leif Egil Loe

Paper Number: SL2012-117

Abstract

Predation risk for animals can be expected to correlate with attributes of habitat structure related to degree of hiding cover. Roe deer (*Capreolus capreolus*) and other forest ungulates may seek hiding cover to avoid detection by predators (for instance human hunters), or choose to avoid hiding cover because it benefits stalking-and-ambush predators (such as the lynx (*Lynx lynx*)). Hiding cover can be point estimated in the field, but completely mapping this quality over a large study area is simply not feasible by ordinary field methods. LiDAR, which measures 3D vegetation structure, is a promising technique for measuring habitat structure relevant for predation risk over large areas. The aim of this study was to predict hiding cover with LiDAR data, using ground truth data from field plots located at known kill-sites and live-sites of roe deer (n=194). The study area is located in Ål and Gol municipalities, Buskerud County, Norway, (8°50'E, 60°40'N). It is a forested valley system with high relief, rising from 200 m a.s.l. at the valley floor up to mountains of alpine character (>900 m a.s.l.) on either side. In and along the valley are roads, towns, a river system, and farmland. A range of habitat types is represented in the area, and the forest is dominated by Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.). As part of a large-scale practical forest inventory going on in the district, LiDAR data were acquired in 2009 and provided wall-to-wall coverage of an area of approximately 872 km^2. The point density was about 1 m^{-2} and first and last returns of the small-footprint laser were used in this study. Field registrations were made in 2011, with plots centered at locations that had been visited by roe deer: 16 hunting kill-sites, 32 lynx kill-sites, 135 live-sites of GPS-collared roe deer, 11 live-sites from feces surveys. A coverboard corresponding to the size and height of a roe deer was used to evaluate hiding cover in the four cardinal directions. The coverboard provides two types of field measures; line-of-sight distance and percentage horizontal cover, the latter measured at 10, 20, 30, 40, and 50m. Differential GPS was used to geolocate the center of the field plots. LiDAR data were exctracted for multiple buffers (Radius of 10, 20, 30, 40, and 50 m) around each of the 194 positions. Height percentiles and density related LiDAR variables were extracted and used as predictor variables in GLMs to predict line-of-sight distance and percentage cover. The paper will report on the model selection procedure, the final models, and the result of leave-one-out cross-validation. This study is the first step towards creating a prediction map of hiding cover for the whole study-area, a map product that can be used in further ecology studies of habitat selection by roe deer facing both human hunting and lynx predation.

Assessing canopy throughfall water flux from airborne laser scanner data and field measurements

Johannes Schumacher[1], Jesper Riis-Christiansen[2]

[1]University of Copenhagen, Forest & Landscape, josc@life.ku.dk
[2] University of Copenhagen, Forest & Landscape, jrc@life.ku.dk

Paper Number: SL2012-119

Abstract

Trees act as an important link in the forest water balance as they effectively intercept, store and transfer water in the canopy-soil interface. Precipitation is intercepted by tree crowns and is evaporated directly back to the atmosphere or is transferred to the ground by canopy throughfall or stemflow. These water fluxes are important as they influence ground water regeneration. However, ground surveys of forest water balance are costly and cover mostly limited plot sizes. At larger scale studies of the forest water balance it is thus desired to integrate forest vegetation composition and retain a detailed physical description of the trees as this has a pronounced impact on the forest water fluxes.

In this study we investigate how canopy structural properties derived from airborne laser scanner (ALS) data are related to canopy throughfall in mono conifer and broadleaf forests. We use ALS data collected for Denmark in 2006/2007 and simultaneous manual canopy throughfall measurements at selected tree species trial sites. The sites we investigate are distributed all over the country and encompass three conifer and five broadleaf tree species. With these data we capture a broad range of growing conditions and the effect of different tree species as they can have different canopy and leaf structures.

From the 3D point cloud of the ALS data we derive height and density metrics. These metrics include height percentiles, maximum and mean height, the proportion of ground hits and the proportion of hits in different height layers of first and last pulse data. These metrics are used to assess the vertical structure of canopies as canopy base height and crown length. Combining the field measurements of canopy throughfall as a proportion of bulk precipitation and independent ALS metrics we derive a linear regression model that best predicts the observations. Our results point to that using simple ALS metrics a realistic estimate of canopy throughfall can be assessed given that the total precipitation is known. These findings open up for more detailed applications in remote sensing and water resource management as the effects of forest types and tree species composition on water balances in a catchment scale context can be derived without costly ground surveys.

Analyzing changes in the forest structural parameters using multi-temporal terrestrial lidar datasets

Shruthi Srinivasan[1], Sorin C. Popescu[1], Ryan D. Sheridan[1] & Nian-Wei Ku[1]

[1] Spatial Science Laboratory, Department of Ecosystem Science and Management, Texas A&M University, 1500 Research Parkway Suite B 217, College Station, TX, 77845, USA and shruthi1389@tamu.edu

Paper Number: #SL2012-120

Abstract

Accurate measures of the forest structural parameters and monitoring their changes over time is essential to forest inventory, managing wildfires and modelling of carbon cycle. Several studies have showed that terrestrial scanning lidar (light detection and ranging) has proved to be accurate in estimating a suite of forest structural parameters such as tree height, crown width, and diameter at breast height (DBH). However, the potential of this technology to monitor forest growth with multi-temporal datasets remains untested. The overall aim of this study was to develop the methodology for analysing multitemporal lidar data sets acquired using a terrestrial lidar scanner for a plot of post oak trees at the Ecosystem Science and Management (ESSM) range area, College Station, Texas. Leaf-on and leaf-off scans were conducted using Leica Scanstation-2 3D laser scanner in 2010 and 2012 respectively, from either sides of the plot of post oak trees. Specific objectives were to: co-register multitemporal scans together; extract individual tree measurements such as tree height and DBH; analyse differences observed for the DBH; analyse differences between scans using a height bins and a voxel method; compare differences in the vertical structure and tree height at plot-level and individual tree level. In additional to comparing the multi-temporal lidar datasets, the method for extracting tree height and DBH was tested on one time terrestrial lidar datasets for nine plots of loblolly pines in Huntsville, TX. Plot-level minimum height, maximum height, mean height, standard deviation and percentiles were calculated and compared between the years. Cylinders were fitted to the 1.2-1.4 m height bin to retrieve DBH. Range rings were created for each stem; and a z polygon was traced to find the highest point, which resulted in the tree height for individual trees. The DBH estimated from lidar data for post oak trees ranged from 0.3-0.6 m, which were consistent with the results from USDA Forest Service. The average tree height decreased from 6.162 m in 2010 to 6.105 m in 2012. Lidar measurements were also validated against the field measurements using regression. The overall results of this study indicate that terrestrial scanning lidar is a promising technology for monitoring forest growth at individual tree level.

Keywords: terrestrial scanning lidar, multi-temporal, co-register, DBH, height bins, tree height

Comparing Forest Inventory and Analysis Plot-Level Biomass Estimates Derived From Airborne Lidar Metrics, National- and Regional-Level Aboveground Biomass Equations, and Forest Vegetation Growth Modeling

Ryan D. Sheridan[1], Sorin Popescu[1] & Demetrios Gatziolis[2]

[1] Spatial Science Laboratory, Department of Ecosystem Science and Management, Texas A&M University, 1500 Research Parkway Suite B 217, College Station, TX, 77845, USA. ryan.sheridan@tamu.edu; s-popescu@tamu.edu

[2] Pacific Northwest Research Station, United States Forest Service, 620 Main Street Suite 400, Portland, OR, 97205, USA. dgatziolis@fs.fed.us

Paper Number: SL2012-122

Abstract

Airborne lidar-derived metrics have been successfully used as surrogates for traditional ground-based forest biophysical measurements, such as aboveground biomass. Typically, height metrics are compared with ground-based estimates of forest aboveground biomass generated using allometric equations. In the United States, both national- and regional-level allometric equations are used to provide estimates of forest aboveground biomass. Since national-level biomass equations are generalized to be equally applicable across the nation, resulting biomass estimates exhibit larger uncertainties than estimates derived from regional-level equations. It is important to quantify the differences between national- and regional-level biomass estimates, and understand how these differences affect regression models used to estimate aboveground biomass from airborne lidar metrics. This study seeks to expand on previous work, which utilized 27 Pacific Northwest Forest Inventory and Analysis Program (FIA) hectare plots (r = 56.42 m) and national-level aboveground biomass equations to estimate plot-level aboveground biomass in eastern Oregon. National- and regional-level biomass equations were used to estimate mean plot-level aboveground biomass (utilizing the four FIA subplots inside each FIA hectare plot) for each of the 27 previously used FIA locations. Airborne lidar metrics and both sets of aboveground biomass estimates were used as inputs for simple linear regression and multiple regression models. The resulting models were compared. The United States Forest Service Forest Vegetation Simulator software (FVS) was utilized to model tree growth at all FIA locations within the study area where temporal data acquisition disparities between the FIA ground data and airborne lidar data were greater than one year. Biomass estimation and comparison of regression modeling results were repeated for plots with modeled tree growth.

Linking Landsat spectral trajectories and LiDAR to characterize forest structure

Oumer S. Ahmed [a,*], Steven E. Franklin [a], Michael A. Wulder [b], Joanne C. White [b]
[a] Geomatics, Remote Sensing and Land Resources Laboratory, Department of Geography, Trent University, Ontario, Canada oumerahmed@trentu.ca, sfranklin@trentu.ca
[b] Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, Victoria, British Columbia, Canada
Mike.Wulder@NRCan-RNCan.gc.ca, Joanne.White@NRCan-RNCan.gc.ca

Paper Number: #SL2012-123

Abstract

In this study we examine forest disturbance over three decades in a coastal temperate forest on Vancouver Island, British Columbia, Canada. We interpret the relationship between LiDAR-derived estimates of canopy cover and the spectral characteristics of forest disturbance trajectories derived from a relatively long time series of 23 Landsat images. The LiDAR data were obtained on a single date in 2004, and are used to estimate forest structural conditions present at the end of the Landsat time series. Through the combination of these datasets we demonstrate: i) the spectral properties of forests and change related to various forest disturbance patterns; and, ii) the efficient characterization of these changes over time. The Landsat images were geometrically and radiometrically normalized and selected spectral indices were computed. A small-footprint LiDAR dataset was used to estimate canopy cover for each Landsat pixel.

First, a regression analysis between LiDAR-derived forest structure and Landsat spectral variables using a forward stepwise method was employed. Landsat-derived Tasseled Cap Angle (TCA) showed the best correlation coefficient ($r = 0.86$) with LiDAR-derived canopy cover. Most possible combinations of the Landsat spectral indices produced significant regression models, with adjusted R^2 ranging from 0.19 to 0.76, and the TCA was used to characterize change in forest structure over the entire image time series using a trajectory-based automated characterization method. When the method was applied to the TCA time series (1972-2011), approximately 71.5% of the study area was found to correspond to 'intact and undisturbed forest'. Areas characterized by disturbance, disturbance followed by revegetation, ongoing revegetation, and revegetation to stable state accounted for approximately 10.2%, 5.3%, 2.2% and 10.5% of the study area, respectively. Finally, we used spectral mixture analysis (SMA) to extract the spectral characteristics of each disturbance class. We compared endmember abundances within and among classes, and these were compared with the LiDAR-derived forest structural estimates. The SMA fraction images offered a clear distinction between disturbance classes in abundances between sunlit canopy, non photosynthetic vegetation (NPV), shade and exposed soil. The LiDAR-derived canopy cover was more strongly correlated with the Landsat derived sunlight canopy endmember fraction ($r = 0.91$) than the LiDAR estimates of mean height ($r = 0.76$). The sunlight canopy fraction tended to decrease as a function of disturbance intensity, and a change in sunlight canopy was explained by a decrease in the amount of LiDAR-derived canopy cover. Overall, this study indicates that moderately accurate predictions of LiDAR-derived canopy cover and height can be obtained using the Landsat-level disturbance class endmember fractions. Further research is needed to utilize these relationships and build on this new capacity to use disturbance class fraction images and LiDAR data to estimate biophysical properties of disturbed forests.

Further Studies of Echidna® Lidar Scanning in California Conifer Stands and New England Hardwood and Softwood Stands

Tian Yao[1], Feng Zhao[2], Alan Strahler[3], Crystal Schaaf[4], Xiaoyuan Yang[5], Zhuosen Wang[6], Zhan Li[7], Curtis Woodcook[8], Darius Culvenor[9], David Jupp[10], Glenn Newnham[11], & Jenny Lovell[12]

[1]Montclair State University, Montclair, NJ, US, yaot@mail.montclair.edu
[2]University of Maryland, College Park, MD, US, fengbjfu@gmail.com
[3]Boston University, Boston, MA, US, alan@bu.edu
[4]University of Massachusetts Boston, Boston, MA, US, schaaf@bu.edu
[5]University of Massachusetts Boston, Boston, MA, US Xiaoyuan.Yang@umb.edu
[6]University of Massachusetts Boston, Boston, MA, US, Zhuosen.Wang@umb.edu
[7]Boston University, Boston, MA, US, zhanli86@bu.edu
[8]Boston University, Boston, MA, US, curtis@bu.edu
[9]CSIRO Land and Water, Melbourne, VIC, AS, Darius.Culvenor@csiro.au
[10]CSIRO Marine and Atmospheric Research, Canberra, ACT, AS, David.Jupp@csiro.au
[11]CSIRO Land and Water, Melbourne, VIC, AS, Glenn.Newnham@csiro.au
[12]CSIRO Marine and Atmospheric Research, Hobart, TAS, AS, Jenny.Lovell@csiro.au

Paper Number: SL2012-125

Abstract

Ongoing work with the Echidna® Validation Instrument (EVI), a full-waveform, ground-based scanning lidar (1064 nm) developed by AStralia's CSIRO and deployed by Boston University in California conifers (2008) and New England hardwood and softwood (conifer) stands (2007, 2009, 2010), confirms the importance of slope correction in forest structural parameter retrieval; detects growth and disturbance over periods of 2–3 years; provides a new way to measure the between-element clumping factor in leaf area index retrieval using lidar range; and retrieves foliage profiles with more lower-canopy detail than a large-footprint aircraft scanner (LVIS), while simulating LVIS foliage profiles accurately from a nadir viewpoint using a 3-D point cloud. *Topographic slope* can induce errors in retrievals of forest structural parameters becASe the horizontal plane of the instrument scan, which is used to identify, measure, and count tree trunks, will intersect trunks below breast height in the uphill direction and above breast height in the downhill direction. A test of three methods at southern Sierra Nevada conifer sites improved the range of correlations of EVI-retrieved structural parameters with field measurements from 0.53–0.68 to 0.85–0.93 for the best method. To test *change detection*, we scanned three New England forest sites at periods of two to three years. A shelterwood stand at the Howland Experimental Forest, Howland, Maine, showed increased mean DBH, above-ground biomass, and leaf area index between 2007 and 2009. Two stands at the Harvard Forest, Petersham, Massachusetts, suffered reduced leaf area index and reduced stem count density as the result of an ice storm that damaged the stands. A new method for retrieval of the forest canopy *between-element clumping index* from angular gaps in hemispherically-projected EVI data traces gaps as they narrow with range from the instrument, thus providing the approximate physical size, rather than angular size, of the gaps. Applied to a range of sites in the southern Sierra Nevada, element clumping index values are lower (more clumping effect) in more open stands, providing improved results. *Foliage profiles* retrieved from EVI scans at five Sierra Nevada sites are closely correlated with those of the airborne Lidar Vegetation Imaging Sensor (LVIS) when averaged over a diameter of 100 m. At smaller diameters, the EVI scans have more detail in lower canopy layers. Foliage profiles derived from processing 3-D site point clouds with a nadir view match the LVIS foliage profiles more closely than EVI profiles in scan mode.

Remote Estimation of Photosynthetic Efficiency Using a Green Terrestrial Laser Scanner

Troy S. Magney[1], Spencer A. Eusden[3,] Jan U.H. Eitel[1,2], Lee A. Vierling[1], Barry A. Logan[3]

[1]University of Idaho, Geospatial Laboratory for Environmental Dynamics, University of Idaho, Moscow, ID 83844, USA. Contact: tmagney@uidaho.edu
[2] McCall Outdoor Science School, University of Idaho, McCall, ID 83638, USA. Contact: jeitel@uidaho.edu
[3] Bowdoin College, Department of Biology, Brunswick, ME 04011, USA. Contact: blogan@bowdoin.edu

Paper Number: SL2012-127

Abstract

The three-dimensional (3D) datasets provided by terrestrial laser scanners (TLS) play an important role in expanding and improving our understanding of ecological processes. Besides providing 3D information, laser return intensity of a green (532 nm) scanning laser may provide information about photosynthetic light use efficiency. During times when leaves are under some environmental stress, excess energy is dissipated via the xanthophyll cycle in order to protect the photosynthetic apparatus of the leaf – a process referred to as non-photochemical quenching (NPQ). Importantly, with decreasing photosynthetic efficiency, the relative concentration of xanthophyll pigments change, which leads to a decreasing reflectance at around 531 nm. However, detecting these changes from passive sensors can be difficult because of variable atmospheric conditions, canopy heterogeneity, and confounding background effects. Here we test if the laser return intensity of a green scanning laser can track changes in xanthophyll pigment activity and NPQ. We imposed four different light intensities on bur oak (*Quercus macrocarpa*), sugar maple (*Acer saccharum)*, and aspen saplings (*Populus tremuloides*), and measured laser return intensity, spectral reflectance, and chlorophyll fluorescence, before taking destructive pigment samples. Our results suggest strong relationships between laser return intensity and NPQ ($0.80 > r^2 > .50$). There was no significant difference ($p<.05$) between the ability of the TLS and the PRI to predict NPQ under all species. Our results suggest that the return intensity of green scanning TLS are suitable for detecting changes in photosynthetic efficiency at the leaf scale. Future research is needed to test if these relationships hold true at the canopy scale.

High spatial resolution automated three dimensional vegetation structure mapping using computer vision

Jonathan P. Dandois[1] & Erle C. Ellis[2]

[1]Ph.D. Candidate, Department of Geography and Environmental Systems, University of Maryland Baltimore County jon.dandois@umbc.edu
[2]Associate Professor Department of Geography and Environmental Systems, University of Maryland Baltimore County ece@umbc.edu

Paper Number: #SL2012-130

Abstract

High spatial resolution measurements of vegetation structure in three-dimensions (3D) facilitate estimates of vegetation biomass, carbon, community structure, fire hazards and other attributes across landscapes for a wide range of management and scientific applications. LIDAR remote sensing (Light Detection and Ranging) from manned aircraft is currently the most popular method for obtaining these measurements. Here we demonstrate a low-altitude unmanned aerial remote sensing system that produces high spatial resolution 3D point clouds similar in quality for vegetation measurements as those obtained from LIDAR, but also including spectral information for each point. This 'Ecosynth' methodology applies photogrammetric 'Structure from Motion' computer vision algorithms to large sets of highly overlapping low altitude (<130m) aerial photographs acquired using off the shelf digital cameras mounted on lightweight (<2 kg) hobbyist-grade unmanned aircraft systems (UAS). 3D point cloud datasets with RGB spectral attributes were produced automatically from digital photographs acquired by a UAS deployed over three 6.25 ha deciduous forest sites in Maryland USA under leaf-on and leaf-off conditions by processing them using a commercial computer vision software package. Point clouds were geocorrected to a horizontal precision of 1.2 - 4.3 m (radial root mean square error; RMSEr) with vertical precisions from 0.4 - 1.3 m RMSE based on ground control point (GCP) marker locations assessed using mapping grade GPS. Understory digital terrain models (DTMs) were generated from computer vision point clouds using an algorithm commonly used for discrete-return LIDAR, with errors ranging from 1.2 - 3.5 m RMSE. Tree canopy height models (CHMs) produced using Ecosynth methods agreed well with field-measured tree heights (R^2 0.64 - 0.78), with similar accuracy as those from LIDAR (R^2 0.61 - 0.77). Plot-level tree height statistics from Ecosynth were more strongly related to LIDAR height estimates ($R^2 > 0.9$) than to field measurements, suggesting that field measurement errors may have been larger than those from point cloud generation and georeferencing. These results demonstrate that Ecosynth methods can serve as an inexpensive and convenient alternative to LIDAR-based methods for acquiring 3D vegetation measurements across landscapes at the 1 km^2 scale and below. This represents a transformative step forward for the remote sensing of vegetation structure, enabling inexpensive mapping on demand by end users of these data, including community foresters and conservation groups. User-deployed 3D vegetation scanning methodologies can also complement existing aerial LIDAR applications, by enabling high-temporal resolution measurements between LIDAR acquisitions and for validation of vegetation structural and spectral traits for future satellite LIDAR/fusion missions.

Characterizing forest succession in a temperate hemlock-hardwood forest type in Upper Michigan via LiDAR remote sensing

Michael Falkowski[1], Robert Froese[1] & Linda Nagel[1]
[1]Michigan Technological University and mjfalkow@mtu.edu

Paper Number: SL2012-132

Abstract

LiDAR (light detection and ranging) is frequently used to characterize forest structural characteristics across a variety of forests types. Accurate quantification of forest structure can aid in operational forest management activities such as silvicultural planning, forest inventory, and wildlife habitat assessment, among others. Of particular interest is employing LiDAR data to classify stages of forest structural development. Indeed, LiDAR data have been successfully employed to classify structural stages in coniferous forests; however, applications in deciduous systems have been limited. The objective of this study was to use discrete return LiDAR to classify forest successional stage across a managed, structurally diverse, hemlock-hardwood forest in the Upper Great Lakes (Upper Peninsula of Michigan) region, U.S.A. Specifically, we focus on using various LiDAR height metrics to classify eight stages of stand development in hemlock-hardwood forests in Upper Michigan. Preliminary findings indicate that LiDAR metrics mirror variation in coincident stand structure and thus can be employed to accurately classify structural stages in hemlock-hardwood forests of the upper Great Lakes region. Since accepted silvicultural practices are often linked to specific stages of stand development, the results presented herein can be leveraged silvicultural planning activities across large spatial extents.

Using Shape-based Metrics to Determine the Boundaries of Lidar Waveforms

Jordan Muss[1], Naikoa Aguilar-Amuchastegui[2] & Geoffrey Henebry[3]

[1]South Dakota State University, jordan.muss@sdstate.edu
[2]World Wildlife Fund US, AguilarAmuchastegui@wwfus.org
[3]South Dakota State University, geoffrey.henebry@sdstate.edu

Paper Number: SL2012-134

Abstract

Lidar systems have evolved over the past three decades from simple rangefinders that measured height profiles from nadir to highly advanced waveform systems that record the heights and intensities of hundreds of returns per pulse as the system scans large swaths of the landscape. While the hardware has improved, the methods used to sample and analyse the data have not kept pace. Common methods such as thresholding and Gaussian decomposition often underestimate the heights of the ground and canopy and fail to delineate the full waveform. These issues can lead to misclassification of the structure of the underlying forest and have been suggested as a leading cause for underestimation of canopy height and other metrics of forest structure by models that rely on lidar data. We propose a recursive method that uses two shape-based metrics, the centroid (*C*) and radius of gyration (*RG*), to identify the lower (ground) and upper (top of canopy) bounds of lidar waveforms. Our algorithm was tested using full waveforms from LVIS data collected in 2005 over the La Selva Biological Station and adjacent private holdings in Costa Rica. The bounds extracted from the lidar data using the recursive algorithm were compared to independent measurements of ground elevation and top of canopy, as well as the estimates of the ground (*Zg*) and top of canopy (*RH100*) elevations reported with the LVIS data.

References

Muss J.D., Aguilar-Amuchastegui, N.,Mladenoff, D.J., and Henebry, G.M. 2013. Analysis of waveform lidar data using shape-based metrics. *IEEE Geoscience and Remote Sensing Letters*, 10(1), 106-110, doi: 10.1109/LGRS.2012.2194472.

Decomposing Forest LiDAR Extinction for Comparison with SAR Interferometric Coherence over the Laurentides Forest, Quebec

[1]Matthew Brolly and [2]Marc Simard

[1] Department of Geographical Sciences, University of Maryland, College Park, MD, USA
[2]Jet Propulsion Lab, California Institute of Technology, Pasadena, CA, USA
Paper Number SL2012-138

Paper Number: SL2012-138

Among remote sensing tools available for forest research purposes the active methods of LiDAR and RaDAR possess the potential to provide faster and more valuable information than many other methods available. One of their major downfalls is the survey cost incurred, particularly in scientific research where many instances require coverage by both sensors, which can prove very costly. Identifying synergies between these methods could potentially reduce the need for forest coverage by both sensors therefore allowing costs to be reduced while crucially providing more information than any one sensor could provide alone. Additional information could even be added to that offered by historical studies.
Here the authors present a comparison of interferometric coherence between measurements provided by radar interferometry, acquired by L-Band UAVSAR, and predictions of this coherence obtained through canopy LiDAR extinction estimates and vertical foliage profile generation made from LVIS waveforms, all acquired from the Laruentides forest, Quebec. These extinction and profile estimates have been obtained from canopy height profiles created from waveforms by accounting for the amount of energy transmitted through to locations deeper in the canopy using the McArthur-Horn transform and the amount of biomass distributed within the vertical profile interpreted by intensity of backscatter. Extinction values are quantified using Legendre Polynomial decomposition and express the potential in forming extinction look up tables for individual forest classes which can be used to better predict forest height metrics using radar interferometry.

Estimation of forest density vertical profiles by modeling occlusion effects in airborne and terrestrial LiDAR data

Nathalie Morin [a] , Richard Fournier [a] , Sylvie Durrieu [b] , Jean-François Côté [c]

[a] Centre d'applications et de recherches en télédétection, Université de Sherbrooke, Quebec, Canada J1K 2R1, Nathalie.M.Morin@Usherbrooke.ca (corresponding author), Richard.Fournier@Usherbrooke.ca
[b] UMR TETIS Irstea-Cirad-AgroParisTech/ENGREF, Maison de la Télédétection en Languedoc-Roussillon 34196 Montpellier Cedex 05, France, sylvie.durrieu@teledetection.fr
[c] Canadian Forest Service, Canadian Wood Fibre Centre, Corner Brook, Newfoundland and Labrador, Canada A2H 6J3, Jean-Francois.Cote@NRCan-RNCan.gc.ca

Paper Number #SL2012-140

Forest density is a key parameter for sustainable forest management and planning of silvicultural operations (e.g. thinning, regeneration), biomass energy estimation, or biodiversity and ecological monitoring. The vertical distribution of vegetation material inside forest canopies is poorly known due to a lack of field measurements, which are difficult to acquire. The development of Airborne Laser Scanning (ALS) provides new metrics for characterizing forest 3D structure. However, ALS has limited capabilities for assessing the understory layer, especially in dense canopies. The ability to identify the vertical profile of vegetation is affected by object occlusion as the ALS signal penetrates deeper into the canopy. Interestingly, Terrestrial Laser Scanning (TLS) provides complementary and detailed information about the vertical distribution of woody and leafy components in the lower canopy. The aim of this study is to compare ALS and TLS vertical profiles of forest density at plot level, in order to quantify the importance of occlusion effect from ALS and to propose a correction model. The study site is in Malahat located in southern Vancouver Island (B.C.), Canada, and is a 0.4 ha retention block of a coniferous stand composed of douglas fir (*Pseudotsuga menziesii*), western red cedar (*Thuja plicata*), and western hemlock (*Tsuga heterophylla*). The methodology can be summarized by a 5-step procedure: (i) extraction of ALS vegetation and single returns probability density function according to height, (ii) use of voxel geometry to transform georeferenced multiple TLS scans into normalized forest density index values, then aggregated into vertical profiles, (iii) comparison of ALS and TLS forest density vertical profiles, (iv) validation with a normalized surface index from a tree architecture model, (v) fitting of a correction model between ALS returns and TLS density voxels. Preliminary results show that ALS and TLS forest density vertical profiles can be approximated by a Weibull function. As expected, the ALS function shows a narrow peak of forest density in the upper part of the canopy. The TLS function reveals a higher degree of skewness and kurtosis, which seems to confirm that TLS detects a higher range of forest density among the height strata, especially in the lower canopy. The fitting of the TLS function to the ALS data allowed us to identify the equation parameters that need to be corrected in order to overcome the occlusion effect in the ALS signal. However, it remains to be determined whether the differences observed between the ALS and TLS are site-dependent or whether they are predictable characteristics specific to each LiDAR system. The applicability of this method to various forest environments, such as deciduous and mixed heterogeneous stands, is currently being investigated.

Keywords: Forest Density, Vertical Profile, LiDAR, Airborne Laser Scanning, Terrestrial Laser Scanning, Voxel

Canary Pine canopy fuel parameters with LiDAR data in Tenerife Island

Alonso-Benito A., Lorenzo-Gil A., Arbelo M., Hernández-Leal P. & Núñez-Casillas L.

Grupo de Observación de la Tierra y la Atmósfera (GOTA). Departamento de Física FEES, Universidad de La Laguna, 38206 La Laguna. Spain aaloben@ull.es

Paper Number: SL2012-142

Abstract

Forest fire manager require accurate maps of the spatial distribution of forest structure, necessary for the execution of simulation models of forest fire behavior. LiDAR data have proven to be an effective tool for the determination of some parameters that characterize the forest structure. In this paper, different regression models were fitted to relate the LiDAR data to a total of six canopy structural variables at the plot level, including basal area (BA), height (H), aboveground biomass (AGB), canopy base height (CBH), canopy bulk density (CBD) and available canopy fuel (ACF). Parameters were estimated from field data obtained in different plots.

The study area, with an extension of 10 km^2, is located in the north of Tenerife Island (Canary Islands, Spain). The altitude ranges from 870 to 1,808 meters above sea level. It includes a very complex geological mixture of terrains with big ravines. The area contains one the most important Canary Pine (*Pinus canariensis* L.) humid forest in the Island, with high influence of trade winds from northwest.

Field data were collected from 50 circular plots with 10 m radius from May to December of 2011. The plots were located every 500 m, forming a grid that covered the whole study area. The coordinates of the center of each plot were recorded with a GeoExplorer XT. The following dimensions (height, crown width, height to crown base, DBH, distance to center of plot and direction to center were measured for each tree within plot.

The LiDAR data were acquired by GRAFCAN with a Leica Geosystems ALS60 since July to August of 2010 from an average altitude of 2,000 m above ground level. The data used in this study were captured in June 2011. The LiDAR system record four returns per laser pulse with 8 bits of intensity register. The FOV is 35-45º, with transversal overlap of 20% and swath width of 1,660 m. Laser point density was of 0.8 per m2. The horizontal and vertical accuracies are 30-50 cm and 15 cm respectively.

Statistical analyses were separately done for height and intensity metrics and for both together. In all cases, there was only a small difference between the models in terms of R^2. The models obtain R^2 values greater than 0.75.

A novel canopy delineation algorithm for discrimination of understory and overlapping crowns with high-resolution LiDAR

Duncanson, L.[1], Dubayah, R.[2], Brolly, M. [3], Cook, B.[4], Parker, G.[5] & Swatantran, A.[6]

[1] Department of Geographical Sciences, University of Maryland, lduncans@umd.edu
[2] Department of Geographical Sciences, University of Maryland, dubayah@umd.edu
[3] Department of Geographical Sciences, University of Maryland, mbrolly@umd.edu
[4] NASA Goddard Space Flight Center, bruce.cook@nasa.gov
[5] Smithsonian Environmental Research Center, parkerg@si.edu
[6] Department of Geographical Sciences, University of Maryland, aswatantran@gmail.com

Paper Number: SL2012-144

Abstract

The ability to derive individual tree information from LiDAR data should greatly improve forest aboveground biomass estimates, and add more detailed information to ecological studies linked to forest structure. Several canopy delineation algorithms have been developed to extract individual tree information from LiDAR point clouds or rasterized Canopy Height Models (CHM). These algorithms generally have difficulty discriminating between overlapping canopies, and fail to detect understory trees. Our novel algorithm uses a watershed-based delineation of a CHM, which is subsequently refined using the LiDAR point cloud. Pseudo waveforms are generated for each preliminary tree polygon, and significant troughs in the pseudo waveforms are assumed to be disparate crowns. The points below a detected trough are extracted from the point cloud and re-processed to generate an understory CHM. Finally, the understory CHM is segmented into understory canopies. Our algorithm is tested using both Echidna and stem mapped field data from the Sierra Nevada, California, and SERC, Maryland. Validation in both a mixed broadleaf system and a conifer system allows for a quantitative comparison of delineation accuracy between the two forest types.

Using data from full-waveform terrestrial laser scanner to estimate above-ground biomass of western junipers

Shrestha, Rupesh[1]; Glenn, Nancy F. [2]; Spaete, Lucas P. [3]; Hardegree, Stuart P. [4]

[1]Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory, Boise, Idaho, USA. shrerupe@isu.edu
[2]Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory, Boise, Idaho, USA. glennanc@isu.edu
[3]Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory, Boise, Idaho, USA. spaeluca@isu.edu
[4]USDA-ARS Northwest Watershed Research Center, Boise, Idaho, USA. stuart.hardegree@ars.usda.gov

Paper Number: #SL2012-145

Abstract

Western juniper (*Juniperus occidentalis* var. *occidentalis* Hook) has been rapidly expanding into shrub steppe and grassland ecosystems in the Intermountain West over the last century, resulting in increased soil erosion, reduced stream flows, and reduced understory plants, among others. Not all effects of woody encroachment are perceived as negative, as woodland development may serve as a carbon sink or as a new source for biofuel production. Estimating tree biomass and understanding how biomass is allocated across different individual tree components provides insights into tree growth and helps explain the uncertainties in biomass estimates. Full-waveform terrestrial laser scanners (TLS) have higher crown penetration than airborne LiDAR or single return TLS, generating dense point clouds that better capture tree components. We present two methods for estimating biomass of tree components using both full-waveform and single-return TLS data. Sixteen juniper trees (height range 1.5-10 m) were randomly selected using a double sampling strategy from different height classes in the Reynolds Creek Experimental Watershed in the Owyhee Mountains, southwestern Idaho, USA. During September-November 2011, each tree was scanned with a full-waveform scanner (Riegl VZ-1000) and a single-return scanner (Leica Scanstation C10). The dry biomass of each component (foliage, branches and main stem) was measured by destructive harvesting of the trees, and allometric relationships were established with TLS-extracted components. The results indicate the waveform, intensity and reflectivity information from TLS improves the ability to extract tree components and hence estimate the component biomass.

Modeling the geometric space of Douglas-fir tree crowns in the northern Rocky Mountains, USA for fire behavior simulation

J. Ferrarese[1], E. Rowell, C. Seielstad

[1]National Center for Landscape Fire Analysis, University of Montana,
jena.ferrarese@firecenter.umt.edu

Paper Number: ### SL2012-146

Abstract

The recent development of physics-based numerical wildland fire simulation models is focusing attention on ways to improve characterization of fine-grained fuel properties in three-dimensions. At tree scale, current studies distribute crown biomass uniformly through simple volumes described by cones and frustums. Fuel characteristics such as crown bulk density are determined from small samples in the laboratory or estimated from allometric biomass equations and are homogenized across the modeled crowns. In this poster, we describe a novel technique for characterizing the denominator of fuel density, (i.e. the geometric space of tree crowns) using terrestrial laser scanner data and apply it to Douglas-fir trees from across the northern Rocky Mountains, USA. The technique exploits the capacity of laser scanning for making many precise measurements, which are used to delineate the occupied crown space. We automate tree selection and canopy delineation in side-lateral scans and percentile-normalize canopy points as a function of height above ground and distance from the center of tree boles. The resulting percentile height-width profiles facilitate direct comparison of trees of varied size and shape and allow for aggregation of data from many trees. Critical percentile hulls are identified and fit with smoothing curves, and crown base heights are objectively determined through analysis of the interaction between percentile crown widths. We compare percentile metrics with actual crown shapes, examine between tree variability, and contrast laser-derived geometry with simple cones and frustums. This study is one component of a larger effort to develop universal models that realistically distribute fuel mass within tree crowns for common conifer species in the northern Rocky Mountains, USA.

A mobile terrestrial laser scanning application for characterizing Mediterranean vegetation structure along forest roads

Maiza N. Dos-Santos[1], Ignacio del Moral, Alexandre Escolà, Ramon Riera-Tatché, Ricardo Sanz, Joan R. Rosell-Polo, Cristina Vega-Garcia.

[1]University of Lleida maizanara@gmail.com

Paper Number: ### SL2012-150

Abstract

Mobile Terrestrial Laser Scanning (MLS) has become a cost effective and useful tool in the acquisition of quantitative data from forests. One of the main advantages that MLS has over the more usual Aerial Laser Scanning (ALS) is its ability to gather data from underneath the canopy at a forest stand level. This approach is essential when analyzing forest structure for many forest management goals, and specifically for linking forest and fire management strategies. Spain, being one of the most affected countries by forest fires in Europe, necessitates the integration of fire prevention actions into forest management, and territorial planning is of utmost importance both economically and ecologically. Fire prevention planning in Spain usually includes fuel treatments in linear structures. "Fuelbreaks" are often designed along forest roads for the purpose of controlling fire size, but also as a human risk reduction strategy. Treatments are extremely expensive, regardless of the technique used (mechanized or manual, prescribed burning, phytocides, grazing), making location and extension of the treatments a decision that needs to be optimized (i.e. stripe width, tree density, shrubs removal). The accurate inventory of vegetation structure along strategic roads is a pre-requisite for evaluating fuel hazard, estimating the costs and improving the decision making allocation process. This inventory is in itself time-consuming and costly, but by using a mobile terrestrial laser scanning along the roads, the process can be made more efficiently.

A MLS system has been evaluated to assess its potential in the acquisition of vertical and horizontal forest structural data, quality and accuracy, over linear tracks. The study area is located in the municipality of Pallejà in Catalonia, Spain. Data were acquired on 16th February 2012 by using a Hokuyo Scanning Laser Range Finder UTM-30LX. The device was mounted in a 4x4 Nissan Pathfinder vehicle, which remained at a reduced constant speed (aprox. 2.5 km/h) during the acquisition of LiDAR point clouds along a forest road. The data were collected from both sides of the forest road simultaneously, being predominantly occupied by Aleppo pine stands (*Pinus halepensis*). This native species is the most extended in Catalonia (20% of the total forest area), and the most affected by fires. The selected site was representative of the hazardous conditions usually found within these forests in the Catalonian region.

Field measurements on individual stems and crowns, understorey height were used for testing quality and accuracy of the parameters provided by the processing of the LiDAR data point cloud. Good results were achieved regarding the characterization of the vertical structure of the forest and in the evaluation of the status of the fuelbreak areas. Under the current high costs of forest fire prevention in Catalonia, we believe these MLS applications can improve current forest inventories, reducing costs and aiding managers in the decision making processes of where to intervene and how to conduct the necessary fuel reduction treatments.

Integration of variable radius plot and lidar data for multi-stand imputation and mapping of forest attributes

Deo, R.K.; Falkowski, M.J.; Hudak, A.; Froese, R.E.

Michigan Technological University, MI-49931, USA

Paper Number SL2012-156

LiDAR remote sensing is a valuable tool for forest inventory and assessment and is beginning to be integrated into operational forest management systems. Techniques that spatially extend traditional forest inventories with LiDAR data typically leverage a fixed radius plot design for the collection of field data. However, numerous efficiencies associated with time and funding could be gained by coupling LiDAR data with variable radius plot (VRP) inventory. The goal of this study is to develop an effective methodology for integrating LiDAR and VRP data to perform a large area forest inventory. This study, which is conducted in the Malheur National Forest in Eastern Oregon, USA, employs a multi-scale, stand-level imputation strategy for generating a spatially explicit forest inventory data. A total of 641 VRPs from 88 stands were used and 37 lidar predictor metrics were extracted at various spatial scales around the VRP centers for model building and identification of optimal lidar grid size for imputation. Accuracy is assessed by comparing the imputed results with independent field data.

Automated water feature delineation and riparian wet areas mapping from lidar data

Michael Chubey[1], Kevin Stehle[1] & Todd Mitchell[2]

[1]Silvatech Consulting Ltd. m.chubey@silvatech.ca
[2]Fugro Consultants Inc. tmitchell@fugro.com

Poster Number: SL2012-162

Abstract

Information about the location and extent of water features is required for many applications including various aspects of environmental planning and management. Accurate generation of water feature information can be costly and time consuming as this often requires production of stereo and/or orthophotos and typically involves labour intensive manual analysis and photo interpretation. This poster will demonstrate two highly automated lidar based approaches for delineation and mapping of water features and riparian wet areas.

In the first example, standing water is detected and delineated using lidar data collected with the FUGRO FLI-MAP system. The FLI-MAP is an airborne sensor designed primarily for collecting lidar data, but contains an additional line scan camera which provides true colour values to each lidar point. An automated object-based analysis procedure was developed to take advantage of unique aspects of the colourized lidar data enabling highly detailed delineation of water features present on the landscape at the time of data collection. The result is a timely, accurate, low cost solution for extracting water features from FLI-MAP colourized lidar data.

In the second example, lidar data were used to produce a riparian wet areas map. In this case, riparian wet areas were modelled using available lidar data in conjunction with known water locations from an existing base features hydropolygon map. A detailed digital elevation model (DEM) was constructed from the original lidar point data. A flow accumulation model was employed to generate a single line water layer connecting surface water polygons. Riparian wet areas were then identified through an automated iterative process where known water locations were used as "seeds" from which the wet areas were "grown" based on DEM derived criteria. The resulting map provides a spatially explicit, detailed, and accurate depiction of critical riparian wet area locations within the project area

The lidar-only approaches demonstrated in this poster eliminate the need for costly orthophotography and feature very short turnaround times due to the almost fully automated nature of the procedures.

Title: LANDSCAPE AND FOREST STRUCTURE AT WOLF-MOOSE KILL AND STARVATION SITES IN ISLE ROYALE NATIONAL PARK. A LiDAR-BASED ASSESSMENT

Luis M. Verissimo[1], Michael J. Falkowski[2], Joseph K. Bump[3]

[1]School of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931, USA. verissim@mtu.edu

[2]School of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931, USA. mjfalkow@mtu.edu

[3]School of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931, USA. jkbump@mtu.edu

Paper ID: SL2012-165

Abstract

Landscape structure and heterogeneity play a potentially important, but little understood role in predator-prey dynamics. For example, prey may strategically take advantage of biophysical structure of their surroundings as a defensive mechanism, and different gender or age classes of prey may be more or less prone to predation in certain landscape conditions.

In this study, we examine whether small-footprint discrete return LiDAR data could be used to assess the influence of landscape and forest structure variables on the location of moose kills by wolves and moose deaths due to starvation. The study covers winter kill sites on Isle Royale National Park over the period 2004-2010, as well as moose malnutrition related death sites from 1996-2010. Wolf-killed moose sites, moose starvation death sites, and random sites are identified and used with a set of Random Forest classification procedures supported by conditional density plot analysis to identify potential predictor variables relevant for each mortality class. Analysis is performed on a island wide approach, as well as in a partition of the island into two strata defined by the winter territories of 3 wolf packs for comparative assessments. The predictor variables, extracted from a set of LiDAR metrics, represent ecological and physiographic characteristics potentially important to kill site locations. These parameters are then integrated in logistic regression models to access, quantify and predict landscape level variables associated with wolf predation of moose as well as infer on starvation site selection factors of moose. Implications for the understanding of top predator-prey interactions at the landscape level, and habitat relations on moose malnutrition mortality sites, are discussed.

Co-registration of ALS data and field plots - sensitivity to plot size and other factors

James W. Flewelling[1]

[1]Seattle Biometrics, Seattle, WA, USA flewelling@seattlebiometrics.com

Paper Number: 169

Abstract

The calibration of ALS-based forest inventory prediction equations is often based on fixed-area plot data. Stem-mapped field plots are typically used for individual crown prediction methods and are sometimes used for area-based methods. In production inventories, plot locations are usually GPS determined, often with positional errors of up to 10 m. Various methods to co-locate the stem map and the ALS data are available. One simple method uses first returns to identify a set of a dispersed tree tops attributed with heights and (x,y) coordinates. An optimization procedure is used to find the presumptive location of the plot center and presumptive error in bearing. The procedure minimizes a weighted score of squared errors in planar location and squared residuals in tree height predictions. The data set used to evaluate positional errors consists of eighty-four cadastral-surveyed 0.08 ha stem-mapped plots with measured tree heights in coniferous and mixed conifer-hardwood stands from Panther Creek, Oregon. Three leaf-on and five leaf-off ALS data sets acquired over a five year period are used for evaluation. Lidar density is about eight first-returns per square meter. Positional accuracy is displayed as a function of simulated plot size up to 0.08 ha and to metrics related to homogeneity of tree top positions within the plots. A second problem is the determination of errors in bearing for terrestrial lidar acquired at a subset of the fixed-area plots. Though the stem-mapped plots and the TLS plots have common monumented center points, the two data sets lacked common references for true north. A simple search algorithm identifies the angular offset that results in the best match between a TLS generated stem map and the field-crew generated stem map.

.

Improving plot-based prediction of live above ground biomass using point-filtered airborne lidar data.

Brian M. Wing[1], Martin W. Ritchie[2], Kevin Boston[3], Warren B. Cohen[4], Alix Gitelman[5]

[1] Oregon State University – College of Forestry brian.wing@oregonstate.edu
[2] US Forest Service – PSW Research Station mritchie@fs.fed.us
[3] Oregon State University – College of Forestry kevin.boston@oregonstate.edu
[4] US Forest Service – PNW Research Station warren.cohen@oregonstate.edu
[5] Oregon State University – Department of Statistics alix.gitelman@oregonstate.edu

Paper Number: SL2012-172

Abstract

Live above ground biomass (AGB) assessments are used in a myriad of forest management applications. Airborne lidar remote sensing has been used successfully to predict live AGB at both the individual tree and the plot level, with the latter being more common in practice. Thus far, most plot-level assessments utilize lidar-derived metrics created using the entire plot's lidar point cloud to predict live AGB. While these metrics have produced promising results, many of the points located within these plot-point clouds are associated with dead trees introducing variability into the modeling framework. In this study, live AGB prediction models are created using three separate sets of plot-level lidar-derived metrics and compared using both model selection statistics and cross-validation statistics. The three sets of lidar-metrics used in the study were; 1) a 'traditional' set created using the entire plot lidar point cloud, 2) a 'live-tree' set created using a lidar point filtering algorithm that attempts to remove points associated with dead trees from the plot point clouds, and 3) a 'vegetation-intensity' set created by filtering and removing lidar points not meeting intensity value thresholds associated with live vegetation. The live-tree lidar metric set produced the best results, reducing prediction variability by 4.3% over the traditional set in plots containing filtered dead tree points. The study highlights the ability of filtering lidar point cloud data to create more useful and accurate explanatory variables for prediction of live AGB. The method should also extend into the prediction of other live tree attributes with similar results.

SilviLaser 2012, Sept. 16-19 September 2012 –Vancouver, Canada

Improving estimation of forest aboveground biomass at plot level using airborne Lidar data to estimate local height heterogeneity

Marc Bouvier [1], Sylvie Durrieu [2], Richard Fournier [3]

[1] UMR TETIS, IRSTEA, 34093 Montpellier, France (marc.bouvier@teledetection.fr)
[2] UMR TETIS, IRSTEA, 34093 Montpellier, France (sylvie.durrieu@teledetection.fr)
[3] Centre d'Application et de Recherches en Télédétection, Depart. of Applied Geomatics, University of Sherbrooke, Quebec, Canada J1K 2R1 (richard.fournier@usherbrooke.ca)

Paper Number: SL2012-175

Abstract

Aboveground biomass (AGB) estimates are required to improve our knowledge on carbon cycle and for ecosystem modelling. Suitable mapping of AGB also supports the implementation of sustainable management strategies and practices that will contribute to forest ecosystem preservation and climate change mitigation. The potential of Lidar to assess AGB at plot level is widely acknowledge [Nelson et al. 1988, Næsset 2004, Van Leeuwen and Nieuwenhuis 2010]. In most studies, biomass estimation are estimated from statistical relationships linking biomass values measured in field inventory, to Lidar metrics extracted from the point cloud data. In general, several height percentiles are selected to describe tree height distribution, and only a few of them remain in the final model. In such approaches, biomass estimations do not take into account horizontal heterogeneity of canopy. The aim of this study is to improve AGB estimation for mono-layer stands by including indicators of the spatial heterogeneity of tree height distribution derived from Airborne Laser Scanning (ALS) data.
As part of the FORESEE project (www.fcba.fr/**foresee**/), a 70 km^2 area covered mainly by Maritime Pine (*Pinus pinaster*) in the Landes forest (South-Western France) was sampled by an ALS system. This ALS has a small footprint and a full-waveform signal. Sixty circular plots (0.1 ha or 0.7 ha each depending on tree heights) were inventoried by traditional field measurements and/or using a terrestrial laser scanner (TLS). Reference biomasses were derived from tree height and diameter at breast height (DBH) measurements using allometric equations.
The current estimates of AGB from statistical relationships do not provide satisfactory results. It is hypothesized that spatial metrics describing local heterogeneity will allow improving AGB in mono-layer stands. Therefore, we investigated field data in order to determine key parameters that could complement those usually derived from Lidar. We identified new parameters allowing to correct the bias due to the extrapolation at plot-level of allometric equations that are valid for individual trees. Combining local stand density with tree mean height was insufficient to correct this bias primarily due to heterogeneity in tree heights. Conversely, adding the kurtosis and the skewness of the tree height distribution to the mean height values turned out useful to correct this bias. Consequently we defined new Lidar metrics aiming at quantifying spatial heterogeneity and skewness of tree height distribution. These metrics were then used to build a new AGB estimation model. The obtained model is compared to biomass estimation calculated from reference measurements. This model reduces error significantly (6%) compared to model based only on mean height. Further studies will be required to investigate the capacity of such kind of model to predict AGB in complex forest stands, especially in multi-layered forests. Full-waveforms data should also be analyzed to improve estimations.

References

Nelson, R., Krabill, W., Tonelli, J., 1988. Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, 24 (2), pp. 247-267.

Næsset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19 (2), pp. 164-179.

Van Leeuwen, M. and M. Nieuwenhuis, 2010. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research* 129(4): 749-770.

Plot age estimation using k-NN and LiDAR

Etienne B. Racine[1], Nicholas C. Coops[2], Jean Bégin[3] & Benoît St-Onge[4]

[1] Laval University etienne.bellemare-racine.1@ulaval.ca
[2] University of British Columbia nicholas.coops@ubc.ca
[3] Laval University jean.Begin@sbf.ulaval.ca
[4] Université du Québec à Montréal st-onge.benoit@uqam.ca

Paper Number: SL2012-181

Abstract

Forest stand age is an important variable to help understand boreal forest dynamics, growth and development. Age is also an important component of site index, since it is used as a reference point to express growth rate in stands. However, age is difficult to measure on the field, making it an expensive variable to obtain. Therefore, height is often used as a proxy for age in many forested ecosystems. Height and age relationship evolution is generally influenced by the underlying relationship with site features such as terrain attributes.

The objective of this study is to improve k-NN age estimation using vegetation structure and site features both derived from LiDAR.

Field data consisted of tree cores from 173 fir-dominated plots collected between 2005 and 2010. LiDAR data was processed to create vegetation and site information for each plot and used as predictors for the k-NN method. Site predictors included elevation, slope and aspect in addition to hydrology and insolation while 95% lidar height, first return penetration rate and parameters from a fitted Weibull distribution were used as vegetation structure predictors.

Age estimation using LiDAR data was highly accurate, with low errors (RMSE < 10 yr). Forest structure variables, derived from LiDAR exhibited the most importance to estimate age, while site predictors such as elevation and slope also improved the estimates, but to a less degree. Our results will allow improvements of forest management by i) lowering the error in stand age estimation and ii) refining the spatial scale at which age is estimated. This fine spatial resolution will allow managers to better understand local forest structure evolution and maintain a diversity of age class distributions.

Single Tree Detection from Airborne Laser Scanning Data Using a Marked Point Process

Junjie Zhang[1] and Gunho Sohn[1]

[1]GeoICT lab, Earth and Space Science and Engineering Department, York Univesity, Toronto, Canada, {jason84, gsohn}@yorku.ca

Paper Number: #SL2012-183

Abstract

There is a clear trend towards individual-tree-based forest inventory. With the increasing availably of high density small-footprint airborne laser scanning (ALS), it has become possible to extract more fundamental tree-based forest parameters, such as tree height, crown size, tree volume and tree species, which are of important economic and environmental concerns from ALS data exclusively.

Single tree detection is usually the basis of tree-wise analysis and parameter derivation and done by segmentation of Canopy Height Model (CHM) recovered from ALS data. Low-level segmentation methods always result in commission and omission errors due to the complexity of forest scenes, which is evidenced by a large body of single tree detection researches. In this research, we presented a probabilistic model for single tree detection based on marked point process aiming to reduce commission error resulted from local maxima based approach. We consider single trees in ALS recovered CHM as a realization of a marked point process of circles. The model takes into account both the geometric characteristics and pair-wise objects interactions of tree crowns in a data term and a contextual term respectively. Different from the conventional marked point process, we proposed to sample our model in a constrained configuration space. The positions of trees are limited to the point set of local maxima extracted from CHM with circular filter window of variable size, and the radius of tree crowns are set as the equivalent radius of corresponding segments of local maxima resulting from marker-controlled watershed segmentation. Finally, the best configuration is sampled by a death process via a steepest decent scheme, through which single trees are detected. Compared with conventional marked point process, our method makes use of priori information can be extracted from CHM and constrains the potential objects within this constrained space, rather than sampling the whole configuration space randomly, which greatly reduces the searching space and accelerates optimization process.

The method is applied on simulated ALS data of forest plots with different degrees of crown overlapping as well as real coniferous forest plots. The tentative detection results show that the proposed method can greatly reduce the commission error and increase the overall detection rate to a large extend, when compared with the traditional local maxima detection based single tree detection method.

Using vehicle mounted LiDAR for Road Safety Inspections on Canada's Resource Roads

David Beleznay[1], Glen Legere[2] & Matt Kurowski[3]

[1] FPInnovations david.beleznay@fpinnovations.ca
[2] FPInnovations glen.legere@fpinnovations.ca
[3] FPInnovations matt.kurowski@fpinnovations.ca

Paper Number: SL2012-184

Abstract

Canada's resource roads are not always engineered to ensure optimal road user safety. In 2007, British Columbia's forest sector recorded 12 fatalities, five of which occurred on resource roads. The following year, the BC government and its partners acted to improve the safety of BC's roads by asking FPInnovations to develop a cost-effective Road Safety Inspection (RSI) service that could prioritize road safety upgrades and traffic control measures. This poster presents the development and implementation of FPInnovations' LiDAR-based RSI service and discusses the service's benefits using results from completed RSIs. Developing the RSI service first involved specifying the requirements for road geometry acquisition and then partnering with companies that operated vehicle mounted LiDAR. Trimble's Trident Analyst software was used to develop a methodology that automates the analysis of the collected point cloud datasets for sight distance calculations and other road parameters. Since 2008, FPInnovations has performed over 800 km of RSIs on BC resource roads, leveraging investments from partners such as Tolko Industries, WorkSafeBC, BC Forest Safety Council, and the BC Ministry of Forests, Lands and Natural Resource Operations. Collaboration with these partners in full-scale operational trials has shown the LiDAR-based RSI service to provide significant cost and time savings compared to traditional RSI methods, and to deliver point cloud datasets that have potential use in value added activities such as as-built modeling and cut and fill calculations.

Comparing the Thematic Accuracy of Imputed Forest Biomass Estimates

Donald Gagliasso[1], Hailemariam Temesgen[2] and Susan Hummel[3]

[1]Graduate Student, College of Forestry, Oregon State University, donald.gagliasso@oregonstate.edu
[2]Associate Professor in Forest Biometrics and Measurements, College of Forestry, Oregon State University, temesgen.hailemariam@oregonstate.edu
[3]Research Forester, United States Department of Agriculture, Pacific Northwest Research Station, shummel@fs.fed.us

Paper/Poster Number: SL2012-186

Abstract

Various methods have been used to estimate the amount of above ground forest biomass across landscapes and to create biomass maps for specific stands or pixel across ownership or project areas. Without an accurate estimation method, land managers might end up with incorrect biomass estimate maps, which could lead them to make poorer decisions in their future management plans.

Previous research has shown that nearest-neighbor imputation methods can accurately estimate forest volume across a landscape by relating variables of interest to ground data, satellite imagery, and light detection and ranging (LiDAR) data. Alternatively, parametric models, such as linear and non-linear regression and geographic weighted regression (GWR), have been used to estimate net primary production and tree diameter.

The goal of this study was to compare various imputation methods to predict forest biomass, at a project planning scale (<20,000 acres) on the Malheur National Forest, located in eastern Oregon, USA. This study compared the accuracy of linear regression, GWR, gradient nearest neighbor (GNN), most similar neighbor (MSN), random forest imputation, and k-nearest neighbor (k-nn) to estimate biomass (tons/acre) and basal area (sq. feet per acre) across 19,000 acres on the Malheur National Forest.

To test these imputation methods a combination of ground inventory plots, LiDAR data, satellite imagery, and climate data were analyzed, and their root mean square error (RMSE) and bias were calculated. For biomass prediction, the k-nn,(k=5) had the lowest RMSE and least amount of bias. The second most accurate method consisted of the k-nn (k=3), followed by the GWR model, and the random forest imputation. The GNN method was the least accurate. For basal area prediction, the GWR model had the lowest RMSE and least amount of bias. The second most accurate method was k-nn (k=5), followed by k-nn (k=3), and the random forest method. The GNN method, again, was the least accurate.

Further analysis is being done to compare the two best performing methods: k-nn (k=5) and GWR, to MSN (the current method used for forest planning purposes on the Malheur National Forest) to examine their predictive abilities at varying geographic scales on the Malheur National Forest.

Tree Volume and Biomass Estimation using Terrestrial Laser Scanner Data: A Case Study of the Mangrove Forests in the Everglades National Park

Emanuelle A. Feliciano[1], Shimon Wdowinski[2] & Matthew D. Potts[3]

[1]University of Miami; efeliciano@rsmas.miami.edu
[2]University of Miami; swdowinski@rsmas.miami.edu
[3]UC Berkeley; mdpotts@berkeley.edu

Paper Number: SL2012-187

Abstract

Wetland forest ecosystems, such as the Everglades National Park (ENP) have greater carbon storage and sequestration capabilities than tropical and temperate forests, especially in their mangrove ecosystems. Mangrove ecosystems form an essential link between terrestrial and aquatic ecosystems providing refuge and food for various animal species and reducing the devastating effects of tropical storms. However, mangrove environments are being threatened by accelerated climate change, sea level rise and coastal projects. Carbon/biomass losses due to natural or human intervention can affect global warming. Thus, it is important to monitor biomass fluctuations in large wetland forests similar to the ENP. Tree volume and tree wood specific density are two important measurements for the estimation of vegetation above-ground biomass (mass = volume * density). Wood specific density is acquired in the laboratory by analyzing stem cores acquired in the field. However, tree volume is a challenging task because trees resemble tapered surfaces. The majority of published studies estimate tree volume and biomass using allometric equations, which describe the size, shape, volume or biomass of a given population of trees. However, the equations can be extremely general and might not give a representative value of volume or biomass for a specific tree species. In order to have precise biomass estimations, other methodologies for tree volume estimation are needed. To overcome this problem, we use a ground-based Light Detection and Ranging (LiDAR) a.k.a Terrestrial Laser Scanner (TLS), which can be used to precisely measure tree volume from its 3-D point cloud. The 3-D point cloud revealed that trees could be modeled as a combination of tapered geometric surfaces called frustums. The volume of the various geometric frustums can be estimated using specific mathematical formulas. We surveyed three mangrove communities: (*Rhizophora mangle*, *Laguncuria racemosa* and *Avicennia germinans*) in three different sites along Shark River Slough (SRS), which is the primary source of water to the ENP. Our sites were located in the Southwestern portion of the SRS that resembles the spatial distribution of mangrove size and shape of the coastal ENP and includes small-, intermediate- and tall- size mangroves. Our ground measurements included both: traditional forestry surveys and state-of-the-art TLS surveys for tree attributes (height, diameter at breast height (DBH), canopy cover) comparison. Traditionally, these attributes are used as input to allometric equations for the estimation of tree volume or biomass. The TLS surveys provided centimeter resolution 3-D point clouds of the below-canopy vegetation. A total of 25 scans were collected in 2011 with a Leica ScanStation C10 TLS, which utilizes a narrow, green (532 nm) laser beam. In this study, we compare our TLS-based tree volume/biomass versus allometric equation-based tree volume/biomass for our vegetation sites. Preliminary results yielded a good agreement between the TLS-based and the allometric-based measurements. This investigation contributes with the Reducing Emissions from Deforestation and Forest Degradation (REDD) project goals, in which large-scale biomass, carbon stock and vegetation structure monitoring are needed.

Automatic Reconstruction of Stem Volume using Terrestrial Lidar

Akira Kato[1], Tatsuaki Kobayashi[2], & L.Monika Moskal[3]

[1]Graduate School of Horticulture, Chiba University, Japan akiran@faculty.chiba-u.jp
[2] Graduate School of Horticulture, Chiba University, Japan ktatsu@faculty.chiba-u.jp
[3]School of Environmental Forest Sciences, College of the Environment, University of Washington, USA lmmoskal@uw.edu

Paper Number: #SL2012-191

Abstract

Laser remote sensing has been used for many applications such as creating DTM, estimating forest biomass, and measuring buildings or artificial objects. Laser can provide high quality 3 dimensional structural data and so it could be used as ground trothing data for satellite remote sensing. In this study, terrestrial lidar was used to estimate the amount of forest biomass as the ground trothing data of satellite remote sensing. Field validation of forest biomass is essential for ensuring the quality of data required to develop both domestic and international environmental policy. Especially, terrestrial laser approach is considered as the best way to measure trees accurately and efficiently while manual measurement likely causes errors.

Aboveground biomass has been estimated from stem volume, which varies depending on tree species. This study aims to automatically quantify stem volume by terrestrial laser to estimate the amount of forest biomass. The terrestrial lidar sensor we used is Riegl VZ400. Digital Terrain Model (DTM) is created from the laser data by the multi-grid method. And watershed segmentation is applied to segment points representing a single tree. Then voxels are created from the segmented points and identify stem parts. Stem points are then extracted and wrapped by the wrapping method using radial basis function and computer graphic techniques to estimate stem volume.

Our study area is located in a commercial plantation area of Japanese cedar and secondary oak forest around Tokyo area in Japan. Stem structure of Japanese cedar is nearly straight, while that of an oak tree is twisted and thus bottom and top of its stem does not match. The method is tested with both coniferous and deciduous tree stems to check its viability. At the same time, 30 destructive samplings were taken to measure the stems directly on the ground. The terrestrial lidar sensor we used is Riegl VZ400.

As a result, 90% of stem from the bottom was captured using vertical scanning and 70% was captured using horizontal scanning. The relative error of stem volume was only 1.3 % and 2.1 % for coniferous and deciduous tree stems respectively. The relative error of conventional stem volume equation is 6.8％ and 12.1% for coniferous and deciduous tree stems respectively. The results show significant improvement in estimating stem volume if terrestrial lidar was used. It is direct, automatic, and non-destructive way to quantify the stem volume. The stem volume equation has been built for many coniferous species since their stems are straight and have the same shape of trunk, which can be quantified statistically. However, the stem volume of deciduous trees like oak cannot be quantified statistically, for its shapes are all different even in the same species. The techniques used for this research is very useful to quantify the stem volume for any shape of stems.

Mapping the forest type for forest management using LiDAR data

Fumihiko Itoh[1], Katsumasa Oono[1], Yoshihira Ogawa[1] & Masako Saitou[2]

[1]Asia Air Survey Co.,Ltd. fmh.itoh@ajiko.co.jp
[2]Kagoshima Prefectural Government Office. masako-saitou@pref.kagoshima.lg.jp

Paper Number: SL2012-196

Abstract

A forest type map is a basic data for forest management. It has been classified by aerial photograph and satellite image. However, there are some places where visual interpretation becomes difficult to classify the forest type by influence of shadow. In addition, it is difficult to identify the difference in composition of tree species in the limited color value of these images. This study introduces new type of image which made from LiDAR data to classify forest types at tree species level.

This study was carried out in national forest of Miyazaki prefecture, Japan, which covers approximately 15 km^2. The study area is covered by broad-leaved secondary forest, plantation of Japanese cedar (*Cryptomeria japonica*), and Japanese cypress (*Chamaecyparis obtusa*) for the most parts. LiDAR data was acquired at 4 points/m^2 density in October, 2011 by ALS50 Ⅱ (Lica Geosystems).

Three types of data are generated from LiDAR data. They are (1) DCHM (Digital Canopy Height Model) image, (2) canopy texture image that is created from DCHM data, and (3) DIM (Digital Intensity Model) image that is derived from intensity data of laser pulse. DCHM image shows tree height. Canopy texture image represents canopy shape and canopy pattern. DIM image indicates intensity of laser pulse reflection. In the field survey, it was found that the difference in the intensity of laser pulse reflection is related to the size of leave area and vertical structure of canopy.

Those images are combined into one image and adjusted color tone. The image named “Laser Forest Type Image”. It can be identified stand height and composition of tree species by color patterns and texture without influence of the shadow.

The forest of study area is classified by visual interpretation using the Laser Forest Type Image, and the result of interpretation is mapped as a forest type map below 7 classes; (1) Broad-leaved forests dominated with pioneer-shrub, (2) Tall secondary forest dominated with wind-dispersed species, (3) Tall secondary forest dominated with high sprouting capacity species, (4) Various species composed mature forest, (5) Plantation of Japanese cedar, (6) Plantation of Japanese cypress, and (7) others (bare ground etc.).

Moreover, it became possible to understand the relation between forest type and terrain by combining Laser Forest Type Map and Terrain Relief Image that is made from DTM (Digital Terrain Model).

Laser Forest Type Image is an effective image which like an aerial photograph and a satellite image for mapping the forest type.

Reference

Kyushu Regional Forest Office, 2012. *“Report on the survey of current forest conditions that utilized remote sensing technology in 2012”*, Kumamoto (Japan), Kyushu Regional Forest Office, Forestry Agency: 73pp.

A comparison between digital terrain models in a forest area of the Island of Tenerife

Lorenzo-Gil A.[1], Isenburg Martin[2], Arbelo M.[1], Alonso-Benito A.[1],

[1]Grupo de Observación de la Tierra y la Atmósfera (GOTA). Departamento de Física FEES, Universidad de La Laguna, 38206 La Laguna. Spain alejandro.lorenzo@ull.es
[2]LAStools developer martin.isenburg@gmail.com

Paper Number: #SL2012-197

Abstract

One of the main products derived from classifying and filtering the LiDAR data are digital terrain models (DTM). Until recently, DTMs were derived from photogrammetric techniques. The purpose of this article is to make a comparison between both types of DTM in a forest area of the Island of Tenerife (Spain).

The study area is located in the northwest of the Island of Tenerife - Spain (28° 25' 32' N, 16º 24' 40' W). It is an area of steep slopes and covered by different plant species: chestnut (*Castanea sativa*), eucalyptus (*Eucaliptus globulus*), pines (*Pinus radiata*, *Pinus canariensis*) and a very common plant association on the island, the fayal-brezal (Myrica faya-Erica arborea).

The LiDAR data was acquired with a discrete return system in summer 2010 with a density of 0.8 points/m2.

The first step was to obtain a DTM derived from LiDAR data using the LAStools software suite. After this, we compared the resulting DTM with an existing photogrammetric DTM and computed a disagreement model. This disagreement model showed elevation differences ranging from 30.32 m to -19.93 m.

To verify which of the two DTM was more in line with reality, we took measures of three elevation profiles using total station techniques,in the areas of greatest disagreement. For each of the three profiles, the results were three lines modeling the ground (topographic measurements-real ground, ground LiDAR, ground photogrammetry).

Analyzing the three profiles demonstrated that LiDAR is able to define in a more realistic manner the ground surface in areas of dense vegetation where the ground is invisible photogrammetric operators.

Our findings are now being used to correct the contour lines in dense forest areas in all the official topographic maps of the Canary island with ground information from the LiDAR DTM.

References

Aguilar, J., Mills, J., Delgado, J., 2009. *Modelling vertical error in LiDAR derived digital elevation models*. ISPRS Journal of Photogrametry and Remote Sensing 65, pp 103-110.

Barreiro, L.,González, E., Corbelle-Rico, E. *Quality control of a digital elevation model generated using LiDAR technology.* FORESTSAT 2010. Operational tools in forestry using remote sensing techniques, pp 93- 96.

Cuenin, R., 1971. *Cartographie Génerale*. Ed. Eyrolles. Paris

Emmanuel P., 1999. *A comparison between photogrametry and laser scanning*. ISPRS Journal of Photogrametry and Remote Sensing 54. pp 83-94.

Flood, M., 2004 *ASPRS Guidelines. Vertical Accuracy Reporting for LiDAR data,* 20 pp.

Höhle, J., Höhle, M., 2009. *Accuracy assessment of digital elevation models by means of robust statistical methods*. ISPRS Journal of Photogrametry and Remote Sensing 64, pp 398-406.

Hyyppä, H., Yu X., Hyyppä, J., 2005. *Factors affecting the quality of DTM generation in forested areas*. Workshop Laser scanning.

Instituto Geográfico Nacional (IGN). Centro Nacional de Información Geográfica (CNIG) 2012a. *Datos geodésicos. Red de nivelación.*

(http://www.ign.es/ign/layoutIn/actividadesGeodesiaRedn.do)

Instituto Geográfico Nacional (IGN). Centro Nacional de Información Geográfica (CNIG), 2012b. *Datos geodésicos. Red de Estaciones Permanentes GNSS.*

(http://www.ign.es/ign/layoutIn/geodesiaEstacionesPermanentes.do

Martin Isenburg, 2010, *LAStools user guides* (http://www.lastools.org/)

Proyecto SATELMAC, 2012. *Mapa de densidad de puntos LiDAR en la isla de Tenerife.*

(http://www.satelmac.com/index.php/es/inicio/132-mapas-de-cobertura-forestal-y-del-terreno-provenientes-de-datos-LiDAR)

Stephen E., Robert, J., Hans-Erik, A., 2003. Accuracy of a high-resolution LiDAR terrain model under a conifer forest canopy. Can. J. Remote Sensing Vol.29, No 5, pp 527-535.

Zhang, J., 2002. *A comparison of digital photogrammetric and LiDAR high resolution digital elevation models*. Research project for the degree of Master of Arts in Geography. West Virginia University.

NEON Airborne Observation Platform Test Flights: Validation of Airborne LiDAR and Hyperspectral Data

Keith Krause[1], Eve-Lyn Hinckley[2], Courtney Meier[3], Dave Barnett[4], Nathan Leisso[5], Thomas Kampe[6], David Tazik[7], Jan van Aardt[8], Crystal Schaaf[9], Alan Strahler[10], Scott Ollinger[11]

[1]National Ecological Observatory Network kkrause@neoninc.org
[2]National Ecological Observatory Network ehinckley@neoninc.org
[3]National Ecological Observatory Network cmeier@neoninc.org
[4]National Ecological Observatory Network dbarnett@neoninc.org
[5]National Ecological Observatory Network nleisso@neoninc.org
[6]National Ecological Observatory Network tkampe@neoninc.org
[7]National Ecological Observatory Network dtazik@neoninc.org
[8]Rochester Institute of Technology vanaardt@cis.rit.edu
[9]University of Massachusetts Boston Crystal.Schaaf@umb.edu
[10]Boston University alan@bu.edu
[11]University of New Hampshire scott.ollinger@unh.edu

Paper Number: # SL2012-199

Abstract

The National Ecological Observatory Network (NEON) is a continental-scale research platform that will collect information on ecosystems across the United States to advance our understanding and ability to forecast environmental change at the continental-scale. The Airborne Observation Platform (AOP) will fly an instrument suite consisting of a visible-to-shortwave infrared imaging spectrometer, a full waveform LiDAR, and a high-resolution digital camera. NEON AOP will focus on several of the terrestrial Essential Climate Variables (ECV) including bioclimate, biodiversity, biogeochemistry, and land use products. These variables are collected throughout a network of 60 sites across the Continental United States, Alaska, Hawaii and Puerto Rico via ground-based and airborne measurements.

A series of AOP test flights were conducted during the first year of NEON construction with the goal to test out instrument functionality and performance, exercise our remote sensing collection protocols, and provide data for algorithm and product validation. These test flights will attempt to address the following questions: What is the optimal remote sensing data collection protocol to meet NEON science requirements? How do aircraft altitude, spatial sampling, spatial resolution, and LiDAR instrument configuration affect data retrievals? What are appropriate algorithms to derive ECVs from AOP data? What methodology should be followed to validate AOP remote sensing products and how should ground truth data be collected? Early test flights around Grand Junction, CO and Ivanpah, CA in May 2012 were focused on radiometric and geometric calibration as well as processing from raw data to Level-1 products. Next, flights were conducted in the Northeast at Harvard and Bartlett Forests in August 2012 with a focus on vegetation chemistry and structure measurements. Vegetation field sampling measurements were performed in coordination with the test flights, including ground LiDAR measurements to validate the airborne LiDAR instrument. An overview of the test flights and ground campaigns will be provided along with any preliminary results that are available.

Using LIDAR to detect fallen trees and wood accumulations in streams: a scaled approach

Mahdi Abalharth

Graduate Student, Department of Geography, University of British Columbia, Vancouver, Canada, Mahdi.Abalharth@geog.ubc.ca

Paper Number: SL2012-203

Abstract

Forests significantly influence watershed hydrology, flow regime and channel morphology, and processes in streams through the abundant supply of wood into the channel network. Such effects are particularly the result of the within-channel accumulation of large woody debris (LWD) and channel-margin trees. Within channel wood has been demonstrated to be critically important in regulating sediment transport and diversifying channel form, thereby also having major effects on aquatic and riparian ecology. LWD exerts its greatest geomorphic influence in channels with physical dimensions similar to or smaller than the size of wood; therefore, wood plays a disproportionately large role in small streams. In these streams, wood controls channel morphology by regulating the temporal, spatial character, and quantity, of sediment stored within the channel zone, and thereby influences channel stability.

Using LiDAR technology, we propose to develop an automated model to identify within-channel wood along channel networks. This model is likely to reduce the need for extensive fieldwork and allow us to identify major locations of wood within the channel network. Although orthophotos can be used to identify logjams, interpretation is a manual process and is less helpful in areas where the canopy can obstruct the view of the channel. In the proposed automated system, we shall examine factors such as what is the smallest channel width in which wood jams can be detected by LiDAR, and the impact of forest canopy on the detection process. Therefore, we will address the following questions in our research: Can LiDAR technology be harnessed to identify wood in channels? How will channel width and canopy view-obstruction impact the identification process?

These questions will be assessed through an integrated program of aerial photograph interpretation, field mapping, and detailed DEM analysis. Study sites will be drawn from areas of the Foothills of Alberta, specifically in Wildhay River and Moberly Creek, where LiDAR-data is available for topographically complex, glaciated drainage basins. To achieve this we propose to conduct a multi-scale integrated field and DEM-based study employing a specific methodological approach, starting our investigation in large streams (>10 m wide) where the canopy does not obstruct the view of the channel. We will use high-resolution digital orthophotos to identify individual logs and logjams. Then, using a combination of 2-D and 3-D analysis of available LiDAR dataset, we will explore various models by which we determine whether a unique signature or pattern of wood jams can be detected in stream channels. Should a reasonable LiDAR-based model be capable of detecting fallen trees in large rivers, the investigation will proceed to smaller streams (3-10 m). The result will also have two scaling components, the minimum size of wood that can be detected in rivers when the view of the channel is unobstructed by the canopy, and how such detection may change as the canopy progressively obstructs the view of channel.

Simulation of lidar waveforms with a time-dependent radiosity algorithm

Huaguo Huang[1,2], Randolph H. Wynne[2]

[1] Key Laboratory for Silviculture and Conservation of Ministry of Education, Beijing Forestry University, Beijing 100083, China (huaguo.huang@gmail.com)
[2] Virginia Polytechnic Institute and State University, Department of Forest Resources and Environmental Conservation, 319 Cheatham Hall (Mail Code: 0324), Blacksburg, VA 24061, USA (wynne@vt.edu)

Paper Number: #SL2012-205

Abstract

Our objective was to assess the effect of multiple scattering on lidar radiative transfer. We have developed a time-dependent radiosity-based model to simulate the propagation of lidar pulses through forest canopies. The model has five distinct steps, as follows: (1) generate a 3D scene using multiscale polyhedra, (2) determine the form/view factors (fraction of energy leaving a given patch that is received at another patch), (3) define sensor position and characteristics, (4) distribute and gather light at each time step (currently 0.5 NS) (5) record the resulting waveform.

The most important model contribution is the explicit characterization of multiple scattering with time. The 3D nature enables simulation of lidar waveforms even with varied topography and clumping vegetation. The incidence angle can be specified. This new model has the potential to provide better approximations of return waveforms. The prototype is being tested using data from the Geoscience Laser Altimeter System (GLAS) acquired over the Eastern Shore of Virginia, USA. Airborne lidar and photos were used to create 3D canopy structures on six one-hectare plots containing GLAS footprints across a range of mean heights. The date differences between GLAS and 3D canopy structures from 2 to 6 years can result in visible tree height growth, leading to a canopy shift on the waveforms. Considering that, model-simulated relative and absolute waveforms show good agreement with measured signals. Results to date indicate that multiply-scattered photons do increase the intensity of the reflected signal, especially the portion originating from the lower portions of the canopy. A peak offset indicator ΔP was defined to assess the multiple effects by fitting two Gaussian functions on single or multiple scattered waveforms.

Sensitivity analysis results show that:
(1) ΔP increases with LAI except several cases when LAI is greater than 5.0;
(2) ΔP increases with the canopy height when other conditions are constant;
(3) ΔP may increase with the slope;
(4) ΔP constantly increases with incident angle from 0.5^0 to 5.0^0 (0.5 to 1 ns);
(5) ΔP decreases when footprint size decreases for sparse canopy or dense canopy (1 ns or 3.5 ns for sparse or dense canopies);
(6) ΔP can be predicted from LAI, H and slope for GLAS:

$$\Delta P=-2.6815+0.1430*H+0.6955*LAI+0.0814*Slope$$ (R^2 =0. 88 and RMSE = 1.4 ns)

The prototype is based on a realistic structural model (RGM); therefore users are able to simulate canopy bidirectional reflectance, light absorption and photosynthesis, as well as lidar waveforms using a same input 3D scene, which provides the possibility to fuse optical and lidar data. Our future work will focus on the data fusion.

Effectiveness of Peat Swamp Forest Protection in Central Kalimantan, Indonesia as Revealed by Multitemporal Airborne Laser Profiling

Yoshiko Maeda[1], Hayato Tsuzuki[2] Limin Swido[3] & Tatsuo Sweda[4]

[1]The United Graduate School of Agricultural Sciences, Ehime University
ttyk-ma151h@hotmail.co.jp
[2]Ehime University zzukky@agr.ehime-u.ac.jp
[3]University of Palangka Raya
[4]Ehime University sweda@agr.ehime-u.ac.jp

Paper Number: #SL2012-206

Abstract

Multitemporal airborne laser profiling of peat swamp forest designated as National park (henceforth protected area) and unprotected peat swamp forest (henceforth unprotected area) in Central Kalimantan, Indonesia in 2007 and 2011 revealed that the unprotected area's vegetation has been degrading relative to the protected area i.e. 80% of the protected area remain the same, 18% forest regrowth and vegetation recovery and 2% vegetation degradation; on the other hand, 74% of the unprotected area remain the same, 20% forest regrowth and vegetation recovery and 6% vegetation degradation. In the unprotected area, the degradation transition from "Low Forest" to "Bush" is majority followed by from "Bush" to "Grassland". There were both along the road, river and the canal which make it easy to get into the forest, the former is due to illegal logging and the latter is due to fire consistent with slash and burn farming. The forest regrowth and vegetation recovery of the both area mainly occurred in burned-out forest where devastated by fire due to the El Nino of 1997, then, there is no difference of vegetation recovery factor between the two sites. Consequently, the peat swamp forest protection in central Kalimantan, Indonesia has been shown to be effective in prevention of vegetation degradation.

Collaborative research to achieve common goals: Calibration and validation of large-area optical remote sensing products using active sensors

Andrew Mellor[12], John Armston[2 3], Andrew Haywood[12], Darius Culvenor[4], Stuart Phinn[2], Peter Scarth[2 3], Richard Lucas[5]

[1] Department of Sustainability and Environment, Melbourne, Victoria, Australia
andrew.mellor@dse.vic.gov.au
[2] Joint Remote Sensing Research Program, School of Geography, Planning and Environmental Management, University of Queensland, Brisbane, Australia
[3] Remote Sensing Centre, Department of Science, Information Technology, Innovation and the Arts, Dutton Park, Queensland, Australia
[4] Commonwealth Scientific and Industrial Research Organisation (CSIRO), Vegetation Sensing & Information - CSIRO Land & Water, Clayton, Victoria, Australia
[5] Institute of Geography and Earth Sciences, Aberystwyth University, Aberystwyth, Ceredigion, Wales, UK

Paper Number: SL2012-210

Abstract

The Joint Remote Sensing Research Program (JRSRP) is a collaborative program that combines research, training expertise and infrastructure to support the work of government land management agencies in Queensland, New South Wales and Victoria. The JRSRP has been established to encourage greater cooperation between the research community and government. Without this cooperation, key research that is common to a number of Australian land management agencies would be both costly and inefficient. Through the JRSRP, a number of research projects are being undertaken to calibrate, validate and improve the accuracy of optical remote sensing derived large-area vegetation information products, using airborne lidar and terrestrial laser scanning. These large-area vegetation products are essential for land management agencies to undertake consistent and regular mapping and monitoring, supporting natural resource management decision making, and meet a range of legislative requirements. In Queensland, field and airborne lidar surveys are being used to independently validate annual Landsat TM/ETM derived foliage projective cover (FPC) – the vertically projected percentage cover of photosynthetic foliage of all strata – across more than 50 sites in eastern Australia. Airborne and ground-based discrete return and waveform lidar are being applied to estimate uncertainty and validate the detection of change in vegetation structure and biomass from mapping being conducted across eastern Australia under the auspices of the Japanese Aerospace Exploration Agency (JAXA) Kyoto and Carbon (K&C) initiative. In Victoria, three 25km^2 instrumented research sites have been established across woodland and open canopy eucalypt forest ecosystems – representative of much of state's 8 million hectares of public forests. Full waveform lidar has been acquired across each site and research fusing-multi source remote sensing data is being undertaken to generate landscape scale forest attribution products across forested ecosystems. In addition, in collaboration with CSIRO, high temporal resolution in-situ autonomous laser (VegNet) scanners are being trialled at these sites to calibrate satellite imagery and improve our understanding of canopy and understorey vegetation dynamics at a very high temporal scale which would be unfeasible using traditional fieldwork methods. Through these research projects, the JRSRP consortium is enabling Australian land management agencies to operationalize the latest research by applying active remote sensing technologies to improve the quality and quantify uncertainty, of satellite imagery-derived large area vegetation products used for mapping, monitoring and reporting.

North West Geomatics Ltd.
and Valtus Imagery Services

Our Solutions are recognized for technological innovation, reliability and our steadfast commitment to delivering high quality geospatial data directly to the desktop.

North West Geomatics Ltd. is a leader in the collection of land resource data with the latest in airborne technology.

Valtus Imagery Services specializes in the storage, integration and distribution of large geospatial datasets through web delivery systems.

North West Geomatics Ltd. | **nwgeo.com**
Valtus Imagery Services | **valtus.com**

North West Geomatics Ltd and Valtus Imagery Services provide end-to-end solutions for the collection, processing, management and delivery of earth information in the form of aerial photography, satellite imagery, elevation information and related data, particularly for large geospatial data users such as resource industries and government organizations.

North West Geomatics Ltd. and Valtus Imagery Services are located at: 212, 5438-11th Street NE Calgary, Alberta. Phone: 403.295.0694